

Une méthode instrumentée pour l'analyse multidimensionnelle des tonalités émotionnelles dans l'interaction

Quignard Matthieu¹, Ursi Biagio², Rossi-Gensane Nathalie³, André Virginie⁴, Baldauf-Quilliatre Heike³, Etienne Carole¹, Plantin Christian¹ & Traverso Véronique¹

¹UMR 5191 ICAR & CNRS

²UMR 5191 ICAR, Université Lumière Lyon 2 & Aix-Marseille Université

³UMR 5191 ICAR & Université Lumière Lyon 2

⁴UMR 7118 ATILF & Université de Lorraine

matthieu.quignard@ens-lyon.fr, biagio.ursi@univ-lyon2.fr, nathalie.rossi-gensane@univ-lyon2.fr, virginie.andre@univ-lorraine.fr, heike.baldaufquilliatre@ens-lyon.fr, carole.etienne@ens-lyon.fr, christian.plantin@univ-lyon2.fr, veronique.traverso@ens-lyon.fr

Résumé. Nous présentons une méthode et un outil exploratoire pour une approche interactionnelle de l'émotion dans les interactions. Cette méthode repose sur le repérage d'indices d'émotionnalité à travers huit dimensions d'analyse : la multimodalité, l'interactivité, les bruits-sons, les marques d'accord et de désaccord, les marques de personnes, les unités macrosyntaxiques, les unités interactionnelles et les répétitions. Deux extraits de corpus provenant de deux équipes différentes sont analysés suivant cette approche. L'outil de visualisation permet d'identifier et de mesurer les phénomènes de concomitance et d'accumulation d'indices à un instant donné, et ainsi d'appréhender la complexité de la construction de l'émotion dans l'interaction.

Abstract. This paper presents a method and an exploration tool for use within an interactional approach to emotion in human interactions. The method is based upon detecting traces of emotionality along eight analytical dimensions: multimodality, interactivity, uttered noises, agreement markers, personal markers, macrosyntactic units, interactional units and lexical repetitions. Two corpus excerpts from different research groups are analyzed following this method. The visualization tool has been designed to support identification of concomitant events and measurement of accumulation of traces at a given point in time, and thus provides better understanding of the complex process of elaboration of emotion during interactions.

1 Introduction

Notre étude propose une méthode et un outil pour appréhender un phénomène complexe, multidimensionnel, telle la construction collective de l'émotion dans les interactions verbales, à travers le faisceau de plusieurs dimensions d'analyse¹. Elle se positionne volontairement en dehors du champ des recherches actuelles sur l'analyse des sentiments, des émotions ou des opinions (voir par exemple le dernier défi « fouille de texte », DEFT 2015²). En effet, notre but n'est pas de caractériser l'émotion ressentie ou exprimée par une personne (Eckman *et al.*, 1969 ; Eckman, 1999), mais d'observer le caractère émotionnel (ou émotionné) d'une séquence verbale et d'analyser par quels mécanismes verbaux, non verbaux et interactionnels ce caractère émerge ou décline. Dans le paradigme de la complexité, l'émotion s'apparente à une propriété du système (à l'échelle macroscopique), plutôt que des individus (échelle microscopique).

Du point de vue méthodologique, le principal biais à éviter est celui qui incite l'analyste à attribuer par *empathie* un caractère émotionnant à un élément de discours. Pour déjouer cette circularité, nous proposons de procéder par des analyses indépendantes de diverses dimensions et de repousser, autant que possible, l'analyse de discours et les propriétés lexicales liées à l'émotion. Un outil a été spécifiquement développé pour rassembler ces analyses et travailler dans un second temps à un niveau macroscopique qui met à distance les individus et la transcription, afin d'étudier de quelle manière dont ces dimensions d'analyse se combinent.

L'étude a été conduite sur un extrait de la base de corpus CLAPI³ (Corpus de L'Angue Parlée en Interaction), puis sur un autre du corpus TCOF⁴ (Traitement de Corpus Oraux en Français), de sorte que sa reproductibilité puisse être observée. Huit dimensions d'analyse ont été explorées.

Nous présentons ci-après les détails de l'approche théorique que nous adoptons, puis la méthode elle-même et ses dimensions d'analyse. Nous détaillons ensuite les principes et la mise en œuvre de l'outil consacré à cette étude. Nous exposons une analyse multidimensionnelle de l'émotion concernant des constructions binaires (mises en évidence par l'approche macrosyntaxique ; Deulofeu, 1977) avec un passage de CLAPI et une analyse émotionnelle sur l'extrait de TCOF, avant d'établir les premières conclusions de cette étude.

2 L'approche interactionnelle de l'émotion

L'émotion peut être approchée de multiples points de vue disciplinaires : philosophie, histoire, sociologie, éthologie, biologie ou psychologie. Les théories biologiques universalistes de l'émotion (Eckman *et al.*, 1969 ; Eckman, 1999) considèrent l'émotion comme une réponse à un événement stimulus, c'est-à-dire comme une réaction organique biologiquement conditionnée par l'occurrence du stimulus et donc universelle. Le lien du stimulus à l'émotion est celui d'une cause à son effet : l'émotion est une réaction réflexe, au sens où il n'est pas possible au sujet de s'y soustraire (par exemple, l'émotion de l'automobiliste qui vient d'échapper à un accident). De ce fait, l'émotion est fondamentalement individuelle ; elle ne peut éventuellement prendre un caractère collectif que dans un processus secondaire dit de socialisation de l'émotion, qui ne touche pas à la nature de l'émotion.

Notre étude s'inscrit dans le cadre de *l'approche interactionnelle de l'émotion* (Plantin *et al.*, 2000 ; Heinemann & Traverso, 2009 ; Peräkylä & Sorjonen, 2012 ; Plantin, 2015a). L'émotion est vue comme un phénomène interactionnel, c'est-à-dire qu'elle est langagière, microsociale, apparaissant au cours de l'action menée par les participants, codéfinie et cogérée par ces participants.

L'émotion est référée à une situation matérielle disruptive, qui peut être partagée par les participants (situation actuelle) ou rapportée par un ou plusieurs participants (situation rapportée). La situation matérielle est dite disruptive dans la mesure où elle ne s'inscrit pas dans le déroulement normal du scénario d'action en cours (*ah merde, décidément je sais pas me servir d'une poêle* (Plantin, 2015a)). Cette situation ne peut être considérée comme un stimulus causal. En effet, dans l'interaction, la situation (actuelle ou rapportée) (1) n'est accessible que par les comportements et les descriptions qu'en donnent les participants. Or, (2) les participants en donnent des descriptions différentes et, en conséquence, (3) ils font émerger de ces descriptions des émotions différentes : une même situation peut être vue par un participant comme justifiant la peur alors que l'autre n'y voit rien de tel :

A : — au fait ça craint la nuit/ qu'elle soit toute seule dans c'te baraque/
L: — non.:

(Plantin, 2012)

Enfin, (4) les descriptions de la situation disruptive et de l'émotion sont corrélées, négociées et évoluent parallèlement au cours de l'interaction.

Il s'ensuit que la notion d'épisode émotionnel linéaire doit être modifiée. L'émotion y est vue comme une excitation affectant un individu, sous l'effet d'un événement stimulus. Cette excitation est décrite comme se développant rapidement jusqu'à un maximum d'expressivité, et décroissant jusqu'à la sortie de l'émotion et la récupération de l'état de calme. Cette représentation, qui présente une certaine adéquation descriptive, a l'inconvénient de « chosifier » l'événement disruptif, de laisser penser que l'émotion est déterminée par cet événement (alors que les deux font l'objet d'un travail de redéfinition dans l'interaction), et que l'émotion se résorbe dans la normalité an-émotionnelle du comportement ordinaire, ce qui n'est vrai que des microémotions.

Schématiquement, nous nous représentons l'interaction, non pas comme se déroulant sur un mode an-émotionnel où surgissent parfois des pics d'émotion, mais comme une courbe où s'enchaînent des moments de plus ou moins grande intensité émotionnelle (voir figure 1). La recherche porte ainsi sur l'affect dans l'interaction, qui peut parfois, mais pas nécessairement, se spécifier en une émotion bien définie (colère, peur, joie...).

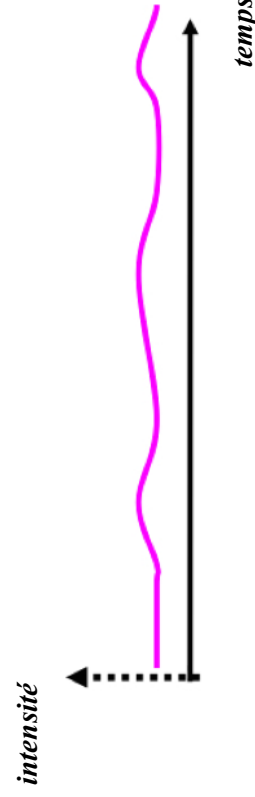


Figure 1. Représentation des fluctuations de l'état émotionnel d'une interaction au cours du temps.

Nous nous proposons ici de saisir ces fluctuations d'affect dans les variations continues ayant une incidence sur les paramètres caractéristiques de l'interaction.

3 Présentation de la méthode

3.1 Le principe de l'analyse multidimensionnelle

Afin de préserver l'analyste du risque d'une empathie inhérente à l'interprétation du discours des participants, nous avons élaboré une méthode d'analyse en deux temps : (1) l'identification d'indices potentiels d'émotionnalité (ou marques *a priori*), puis (2) l'analyse d'un phénomène particulier sur la base des indices observés *a priori*. Nous présentons ci-dessous les détails de la première étape.

Nous adoptons une démarche couplant une approche empirique avec les résultats d'études antérieures qui ont établi des inventaires exhaustifs des marqueurs linguistiques de l'émotion (Kerbrat-Orecchioni, 2000 ; Plantin, 2011). Nous suggérons de considérer comme *a priori* pertinentes un certain nombre de dimensions d'analyse, dont la pertinence réelle ne sera évaluée qu'*a posteriori*. Cette posture exploratoire nous impose en contrepartie de proposer un nombre de dimensions suffisamment important pour que des phénomènes émergents puissent être observés et (in)validés.

Comme annoncé *supra*, nous avons retenu huit dimensions que nous détaillons ci-après. Nous avons opté pour le terme de dimension, et non de niveau ou strate, pour ne pas introduire de rapport d'ordre ou de dominance de l'un(e) sur l'autre. Nous choisissons de mettre l'accent sur les marques verbales ne portant pas en elles-mêmes un sémantisme lié à l'émotion (marqueurs de discours, marqueurs de personnes), sur les structures macrosyntaxiques et interactionnelles, sur les segments répétés ainsi que sur les dimensions non verbales ou comportementales (l'émotion est portée par le corps, dont la voix, et pas seulement par les mots).

L'extrait choisi pour l'expérience, « Jaune Fluo », est une séquence du corpus « Repas, conversations entre étudiants » disponible dans la base CLAPI⁵. Il s'agit d'une conversation de la restitution d'un travail remis dans le cadre de ses études, l'enseignante l'ayant critiquée devant ses camarades (récit de la honte).

L'annotation est réalisée avec l'outil Praat (Boersma & Weenink, 2015) ; chaque dimension a nécessité une ou deux lignes d'annotation.

3.2 Dimension 1 : la multimodalité

La dimension multimodale a pour vocation de mettre en évidence les signaux corporels émis par les participants en rapport à leur état affectif (gestualité, regards, etc.). Néanmoins, nous n'avons exploité dans cette étude que le versant audio de la séquence filmée mentionnée ci-dessus et nous sommes penchés sur les éléments de la matière sonore qui sont susceptibles d'être reliés à la tonalité émotionnelle. À cet égard, nous avons considéré les productions verbales caractérisées par un débit de parole spécifique, qui représente une altération perceptible du débit normal. Ainsi, les traits dénotant un débit plus rapide (ou accéléré (ACC)) ou un débit plus lent (LENT) ont été introduits.

La caractérisation perceptuelle des syllabes a été notée par le biais de deux traits : l'un concernant la durée, et plus précisément la production d'un allongement (ALL), l'autre portant sur l'intensité. Ce dernier trait signale les segments caractérisés par une saillance perceptuelle⁶ (S).

En ce qui concerne l'articulation des annotations, la saillance a été notée séparément car elle concerne parfois un segment qui est déjà caractérisé par un autre trait multimodal (ACC/LENT ou ALL).

3.3 Dimension 2 : l'interactivité

La dimension « interactivité » rassemble les traits propres à l'interaction pouvant affecter la tonalité émotionnelle de celle-ci. Ces traits concernent l'organisation des prises de parole des locuteurs : les pauses (P) pouvant marquer des apaisements ou, au contraire, un embarras, les chevauchements de parole (O), l'alternance simple des tours (C), voire l'enchaînement immédiat (sans pause (E)) entre deux locuteurs différents indiquant une certaine vitalité de l'interaction.

On notera que les chevauchements et les pauses sont des annotations portant sur des segments, tandis que les traits marquant l'enchaînement des tours concernent les frontières de tours (des *points* au sens de Praat, auxquels on attribue une durée conventionnelle ; voir *infra*, note 11). Les annotations de cette dimension découlent exclusivement des changements de locuteurs et ne dépendent aucunement de ce qu'ils disent ou émettent dans le tour (mots, rires ou autres productions vocales). Elles peuvent être produites automatiquement par un outil de reconnaissance de locuteurs⁷.

3.4 Dimension 3 : les « bruits-sons »

Cette dimension répertorie les bruits et vocalisations non lexicalisés, c'est-à-dire les sons qui n'entrent pas dans la liste des interjections. La frontière entre bruits et interjections est néanmoins floue et le degré de lexicalisation ou de conventionnalisation doit être interrogé au cas par cas. Si l'interjection est généralement considérée comme langagière bien que mettant à rude épreuve les règles du système, les bruits et vocalisations sont d'habitude délaissés par la linguistique qui les voit comme ne faisant pas partie du système. Ils sont analysés en tant qu'*expression primaire* (à la suite de Wundt, 1911) ou selon leurs fonctions dans l'interaction (dans l'approche de la linguistique interactionnelle). Ces fonctions, qui ont été décrites pour certaines de façon détaillée, peuvent être très variées (voir Tolins, 2013 ; Hoey, 2014 ; Holt, 2011 ; Hepburn & Potter, 2012). Pour cette étude, nous avons choisi de classer comme bruits-sons toutes les formes non lexicalisées, indépendamment de leur fonction dans l'interaction.

Dans cette dimension, nous séparons les bruits classiques et bien étudiés dans l'analyse interactionnelle, tels que les Rires (R) et les Souffles (S), d'autres bruits-sons dont le fonctionnement est encore beaucoup moins connu. Dans ce deuxième groupe, nous distinguons les vocalisations à proprement parler (V) – c'est-à-dire des segments voisés (« wa:: », « o:w ») – des bruits sourds (B), comme les raclements de gorge, « ts », « ch ch ch », etc.

3.5 Dimension 4 : les marques d'accord et de désaccord

Les marqueurs discursifs⁸ jouent un rôle fondamental dans l'interaction par la flexibilité et la variabilité de leurs usages. L'annotation consiste à repérer ces marqueurs tout en distinguant ceux qui indiquent en contexte une convergence de ceux qui – toujours en contexte – signalent une divergence⁹. Ainsi, un même marqueur n'est pas intrinsèquement positif ou négatif, mais porte cette polarité de manière afférente (voir, par exemple, Rastier, 1987).

3.6 Dimension 5 : les marques de personnes

Les marques de première et seconde personne (pronoms personnels, déterminants possessifs, flexions de verbes, etc.) sont les marques naturelles de l'interlocution et du positionnement du locuteur vis-à-vis de son interlocuteur. Elles permettent d'adresser la parole et d'opposer des discours. Dans certains cas, elles peuvent à elles seules mettre en évidence des conflits (Denis *et al.*, 2012) ou la polémique comme dans les débats télévisés entre les deux tours des présidentielles (Kerbrat-Orecchioni, 2012). L'annotation est très simple : on appose l'étiquette « Je » sur toutes les formes portant la marque de la première personne (*je, moi, mon, ma, mes*, ainsi que tous les verbes conjugués à la première personne du singulier), tandis que l'étiquette « Tu » est apposée sur les formes à la deuxième personne du singulier ou du pluriel (*vous* de politesse). Cette annotation peut être très facilement réalisée par un annotateur morphosyntaxique.

3.7 Dimension 6 : la macrosyntaxe

Rappelons que la macrosyntaxe est définie comme une description de l'au-delà de la rection, par opposition à la microsyntaxe, syntaxe de rection¹⁰, dite encore, simplement, syntaxe. Pour Blanche-Benveniste *et al.* (1990 : 113), à l'origine du cadre théorique aixois de micro et macrosyntaxe sur lequel nous nous appuyons, « il s'agit des relations qu'on ne peut décrire à partir des rections de catégories grammaticales ». Dans ce cadre théorique, l'unité minimale centrale de la macrosyntaxe est le noyau, capable de « former un énoncé autonome » (Blanche-Benveniste *et al.*, 1990 : 114). Dans le prolongement de travaux récents (au sein, notamment, du projet ANR Rhapsodie (voir, par exemple, Lacheret *et al.*, 2011)), les éléments détachés à gauche sont analysés comme des prénoyaux et les éléments détachés à droite comme des postnoyaux, tandis que ceux insérés au milieu du noyau sont appelés innoyaux. Cette dernière catégorie, toutefois, a été envisagée ici de manière large : elle inclut des éléments non régis qui, de même que les prénoyaux et les postnoyaux, ne portent pas de modalités, mais aussi des parenthèses non précédées d'un subordonnant et comportant un verbe fini. Il convient en outre de souligner que les marqueurs ont été considérés comme ne pouvant constituer en eux-mêmes des prénoyaux ou des postnoyaux (à la différence de ce qui se produit, par exemple, dans le cadre théorique représenté dans Morel & Danon-Boileau (1998), où des éléments tels que *et* et *bon*, « ligateurs » respectivement « discursif » et « énonciatif », relèvent *per se* du champ gauche appelé « préambule »).

3.8 Dimension 7 : les unités de construction du tour

Les unités de construction du tour (en anglais *Turn-Constructional Units*, *TCU*) sont les composants élémentaires du tour de parole. Dans le déroulement de la parole, ces unités sont délimitées par des points ou des places où une transition est possible. À ces emplacements, un changement de locuteur peut se produire ou pas, s'il s'agit d'un tour multi-unités où le même locuteur poursuit. La particularité de cette approche de la parole en interaction est de considérer les unités comme en train d'être construites collaborativement par les différents participants (voir Schegloff, 2007). Les marques utilisées par les locuteurs pour indiquer les places de transition possible sont notamment d'ordres syntaxique, prosodique et pragmatique. L'ensemble fonctionne sur la base de deux principes de production : la projection (un élément à l'instant *t* projette un élément suivant, par exemple dans une construction syntaxique) et l'incrémentation (le rajout au fur et à mesure, voire après coup, d'éléments additionnels à une construction).

Dans nos annotations, nous avons dégagé trois types d'unités :

- les TCU simples, qui peuvent éventuellement faire partie de tours multi-unités ;
- les TCU qui construisent des listes ;
- les TCU abandonnées.

L'exemple ci-dessous montre qu'une unité interactionnelle peut éventuellement courir sur plusieurs tours de parole :

142	M	ben il faut vraiment trouver la:	[la la	
143	J		[-<(accélééré) et puis] faut qu'>	
144	M	le TON qui fait que <(accélééré) ça se voit pas trop non plus quoi		
145		et que ça>		

M est en train de produire une TCU, l'allongement du déterminant « la: » (ligne 142) projette un élément nominal du syntagme commencé. Le tour de J (ligne 143) est produit en chevauchement. Il représente un tour abandonné intercalaire (une TCU abrégée). Ensuite, M poursuit la construction de son énoncé par une autoréparation (« la: la la » ligne 142, « le TON qui fait que » ligne 144), insérée dans la même unité interactionnelle précédemment chevauchée par J. Il en résulte un tour étendu comportant une seule TCU.

3.9 Dimension 8 : les répétitions

Le phénomène de répétition a été étudié en tant que ressource utilisée par les participants durant l'interaction, qu'il s'agisse d'une répétition par le même locuteur dans le même tour de parole ou dans des tours successifs (autorépétition), ou de la reprise d'un précédent énoncé par un autre locuteur (hétérorépétition ; Groupe ICOR, à paraître). Les segments peuvent être répétés à l'identique ou avec modification (Couper-Kuhlen, 1996). Seules les répétitions lexicales sont prises en compte ; les faux départs et les marqueurs sont exclus.

La convention d'annotation adoptée est R0 pour le segment initial, Rn pour le segment répété à l'identique la *n*-ième fois, Rn+ pour le segment répété avec addition (*c'est possible => c'est tout à fait possible*) et Rn- pour le segment partiellement répété avec soustraction (*elle envoie pas mal de mails groupés => elle envoie un mail*).

4 Un outil pour l'exploration de la multidimensionnalité

À la fin de cette première étape d'analyse, nous disposons d'une transcription richement annotée puisque huit grandes catégories d'indices y figurent. Or, ce matériau s'avère particulièrement difficile à manipuler par la suite étant donné que :

- le niveau verbal est toujours présent, et offre toujours autant prise à l'empathie ;
- les annotations surchargent la lecture alors qu'il n'est pas garanti qu'elles soient toutes pertinentes ;
- l'appréhension de la dimension temporelle est difficile tant dans la transcription en liste (la longueur du texte n'est pas une mesure sérieuse de la durée des tours, notamment en raison des pauses) que pour l'annotation dans Praat qui se prête assez peu à une vision d'ensemble.

Nous avons donc pris la décision de concevoir un outil de visualisation interactive de notre corpus annoté, dont les détails sont donnés ci-après.

4.1 Les objectifs de l'outil

Pour répondre à nos objectifs, l'outil doit remplir les fonctions suivantes :

1. proposer une représentation graphique des indices choisis (cf. § 3) permettant de ne plus afficher le verbal et d'ainsi de se dégager de l'inévitable emprise de l'empathie dans l'interprétation des résultats ;
2. présenter les indices dans leur temporalité pour rendre compte de leur présence, de leur absence et de leur durée mais, également, pour mettre en évidence des concomitances ou des faisceaux d'indices ;
3. permettre d'explorer interactivement les effets de chacune des dimensions pour dégager les indices les plus pertinents ;
4. rendre possible la modélisation d'une propriété (dans notre étude, l'émotion) d'un moment de l'interaction émergeant par *accumulation* d'indices. Cet outil n'a pas vocation à produire des annotations automatiques des émotions (à l'instar des travaux en traitement automatique des langues qui s'appuient sur une typologie d'émotions ou de sentiments ; Hamon *et al.*, 2015), mais à visualiser un ensemble d'annotations déjà réalisées pour en faciliter l'analyse et l'interprétation.

L'outil propose pour le point 4 (c'est-à-dire la modélisation de l'émotion par accumulation d'indices) un modèle simple et exploratoire :

1. Tous les indices sont égaux sans aucun ordre ni aucune hiérarchie. On ne présume pas qu'un indice ou qu'une dimension soit *a priori* plus fort, plus « impactant » qu'un autre.
2. Tous les indices sont positifs. On n'intègre pas actuellement d'indice à impact négatif, pouvant infléchir la courbe vers le bas.
3. L'impact est lié au segment qui le porte et cesse lorsque le segment porteur arrive à son terme¹¹.
4. Il n'y a pas d'« effet de séquence » : le fait qu'il y ait une pause avant un rire ne confère pas plus de poids (ou d'effet) au rire (ou à la pause).

5. L'interaction est d'autant plus émotivée que les indices sont concomitants. Cette concomitance peut être due à plusieurs segments verbaux émis par plusieurs locuteurs en chevauchement, ou à un seul segment verbal d'un même locuteur présentant des indices annotés dans plusieurs dimensions. Par exemple, un marqueur discursif prononcé de manière saillante impacte doublement l'interaction.
6. Tous les indices ne sont pas nécessairement pertinents. Nous avons élaboré une liste étendue d'indices pouvant affecter l'émotivité. Il est possible que certains indices soient opérants dans certaines situations et ne le soient pas dans d'autres.
7. L'outil calcule le niveau moyen (ou la densité moyenne d'indices), ce qui permet de savoir à chaque instant si la valeur observée est élevée ou basse par rapport au reste de l'interaction.

4.2 Mise en œuvre technique

Le choix du format graphique s'est porté sur SVG car il s'agit d'un format standard XML aisé à produire et tout aussi aisé à manipuler par les technologies web (HTML, CSS, Javascript). Dès lors, l'outil se réduit à un ensemble de scripts et de feuilles de style embarqués dans un navigateur ; il est donc parfaitement portable.

La métaphore graphique est réalisée par une feuille de style XSL qui dispose chaque *tier* de Praat en ligne ; chaque segment Praat devient un rectangle ; chaque point (ou frontière) est matérialisé par un cercle. La largeur du rectangle est proportionnelle à la durée du segment. Une couleur est associée à chaque dimension au moyen d'une feuille de style. Nous obtenons ainsi la frise suivante (Figure 2), où le temps est représenté en abscisse et où les dimensions sont affichées les unes au-dessus des autres :

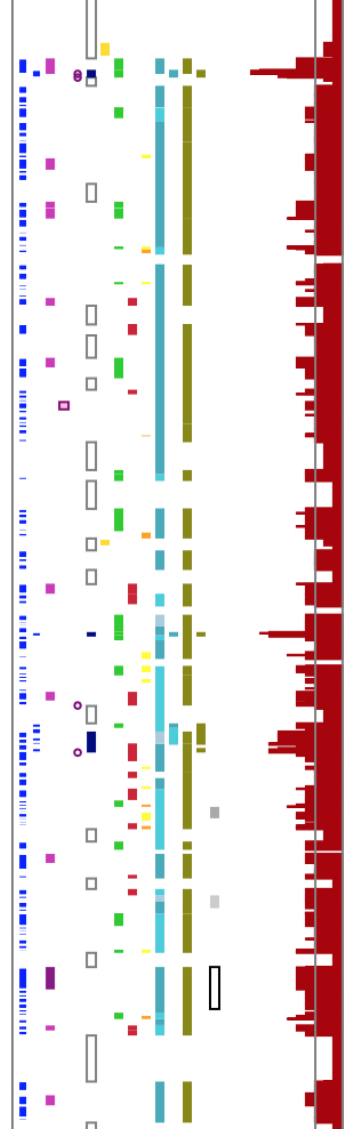


Figure 2. Frise représentant les annotations dans le temps et la fonction d'accumulation (en rouge).

En bas de la frise, en rouge, l'outil construit une courbe correspondant à l'accumulation d'indices. La ligne horizontale grise figure le niveau moyen.

Une librairie JavaScript a été conçue pour gérer la partie interactive de l'outil. Ce programme permet de recalculer l'accumulation d'indices en temps réel, d'effectuer des zooms sur la frise, de jouer l'extrait audio au voisinage du curseur et d'afficher la transcription dans ce voisinage. Cet outil peut donc aussi permettre de (re)vérifier les annotations.

Ci-dessous (Figure 3), le panneau de contrôle interactif de la frise et l'affichage de la transcription deux secondes autour du point courant :

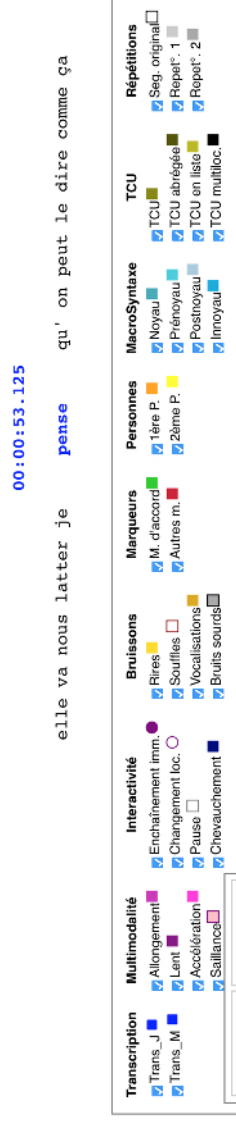


Figure 3. Panneau de contrôle permettant de choisir les marqueurs à intégrer dans la fonction d'accumulation.

4.3 Déploiement de l'outil sur une transcription

Étant donné une transcription dans un format quelconque et son fichier audio, il s'agit dans un premier temps de produire un fichier Praat avec un alignement au mot (ce sont en effet les mots qui portent les annotations). Nous préconisons pour cette tâche l'outil JTrans¹², développé au LORIA par l'équipe Synalp. L'annotation multidimensionnelle est à poursuivre dans Praat en respectant l'ordre des dimensions (pour l'ordre des *tiers*) et la codification correspondant à chacun des indices. Une fois que la phase d'annotation alignée est terminée, on transforme le fichier Praat en fichier XML/SVG à l'aide de la feuille de style XSL *ad hoc*.

On note que cette technique de visualisation interactive peut être appliquée à n'importe quelle transcription alignée, réalisée dans le logiciel Praat ou un logiciel analogue. Il suffit de disposer d'un certain nombre de couches d'annotation et de segments alignés dans le temps.

5 Les analyses

Dans le cadre de l'étude présentée ici, nous avons tout d'abord sélectionné un passage d'un corpus de la base CLAPI afin de tester notre méthode instrumentée pour l'analyse des tonalités émotionnelles dans l'interaction, avant de mettre à l'épreuve cette méthode sur un extrait d'un autre corpus, TCOF, et d'ainsi tester sa reproductibilité.

5.1 Analyses de trois extraits de la base CLAPI

Pour les analyses d'extraits de la base CLAPI, nous avons choisi de nous pencher sur une sorte très particulière de constructions binaires (Deulofeu, 1977), que nous présentons ci-après.

5.1.1 Les constructions binaires

Rappelons que les constructions binaires au sens large consistent en « deux éléments d'énoncé dans une relation de couple, toujours réalisée sur une chaîne rompue [...] et non marquée par des catégories grammaticales » (Deulofeu, 1977 : 37). Au sein de ces constructions, nous nous intéressons à ce que Deulofeu a considéré comme « la forme la plus pure de la relation binaire » (1977 : 50 ; voir aussi Astésano *et al.*, 2008) et dont il a donné pour exemple (parmi d'autres) : *L'Éducation Nationale les cantines se payent au trimestre*. Dans de tels cas, on a affaire à un détachement sans rappel : le syntagme nominal n'est pas repris anaphoriquement dans le segment suivant. En outre, il n'existe pas de relation syntactique, de lien rectionnel, entre l'élément détaché et le segment subséquent. Celui-là ne peut non plus être déplacé (**Les cantines se payent au trimestre l'Éducation Nationale*). Deulofeu précise que « [l]a relation découle d'une intégration marquée par l'intonation » (1977 : 50). Sur un plan pragmatique, Prévost propose, pour les détachements de cette sorte, d'analyser l'élément détaché, non comme

un topique, mais comme un cadre « dans lequel la relation prédicative est pertinente » (2003 : 121-122). De telles constructions nous semblent se prêter particulièrement bien à une analyse multidimensionnelle.

En analyse conversationnelle, les constructions binaires – qui ne sont jamais thématisées en tant que telles, les constructions syntaxiques *per se* n'étant pas l'objet des analyses – entreraient dans la catégorie des tours multi-unités, décrits par le phénomène de *projection*, qui fonctionne de la manière suivante :

- sur le plan syntaxique, une première TCU en projette une suivante pour des raisons syntaxiques, ex. « tous ceux qui sont blancs/ » projette une suite « j` comprends pas » ;
- sur le plan prosodique, une première TCU porte une intonation non finale, ex. « tous ceux qui sont blancs/ », laissant attendre une suite « j` comprends pas ».

Dans notre passage, nous avons sélectionné trois constructions binaires de cette catégorie. Nous en analysons deux en détail, en adoptant pour chacune d`abord une approche macrosyntaxique, puis une approche interactionnelle.

5.1.2 Extrait 1

La convention de transcription utilisée est la convention ICOR¹³. Les symboles • indiquent les frontières de TCU.

1	J	•h t` auras ja:mais cinq\•
2		(0.5)
3	M	•ah si\ avec ma prof c`est possible hein • (0.5) • parce que\ e- • je sais/
4		qu`elle nous attend avec une batte de base ball derrière la
5		portE le jour d` la rentrée hein •
6		(0.4)
7	J	•ah là là:: •
8		(0.6)
9	M	•c`est tout à fait possible •
10		(2.5)
11	-> M	•une interro qui était censée êt` facile/ • ↓la: moyenne de la
12		classe^ • ↓sept •
13		(0.2)
14	J	•TSK •OUAIS::// •Hu.hhu facile hein/ •
15		(1.3)

Analyse macrosyntaxique

Dans l`approche macrosyntaxique, dans la ligne 11, on compte deux prénoyaux, *une interro qui était censée être facile* et *la moyenne de la classe*, suivis du noyau *sept*. Le premier prénoyau (mais non le second) participe d`une construction binaire au sens strict où nous l`entendons ici. On peut parler de détachement sans rappel dans le segment subséquent, lequel est constitué du second prénoyau et du noyau : le syntagme nominal *une interro qui était censée être facile* n`y est pas repris anaphoriquement. On rappelle que ce premier segment n`est pas déplaçable : **la moyenne de la classe sept une interro qui était censée êt` facile*. Les constructions binaires au sens strict sont à distinguer des structures équationnelles, dont participent par exemple le second prénoyau et le noyau dans *la moyenne de la classe, sept*. Dans ce dernier cas, une paraphrase de la sorte *la moyenne de la classe est sept* et une permutation telle *sept, la moyenne de la classe*, où le prénoyau devient postnoyau, sont possibles. Enfin, on peut noter une certaine complexité de cette construction binaire, dans la mesure où le deuxième élément de la structure n`est pas un simple noyau, mais un ensemble constitué du second prénoyau et du noyau.

On remarquera pour cet exemple que, sur un plan sémantique, la construction binaire repose sur les relations lexicales unissant d'une part *interro* et d'autre part *moyenne*, *classe* et *sept*, ce dernier terme, évaluatif, se chargeant, de par le contexte, d'une valeur axiologique.

Analyse interactionnelle

Dans l'approche interactionnelle, on notera pour commencer que la construction binaire intervient dans le contexte d'un développement construit autour de l'affirmation *avec ma prof c'est tout à fait possible* (ligne 3, repris ligne 8). Le tour multi-unités arrive comme un argument soutenant cette affirmation. Il est composé de trois TCU qui ne sont pas séparées par des pauses, mais qui sont nettement contrastées par des sauts mélodiques – les TCU 2 et 3 commençant beaucoup plus bas que la fin de la TCU à laquelle elles font suite – et par des changements d'intensité (montée de l'intensité sur *sept* après une baisse notable). L'unité 1 a une intonation montante, qui projette clairement une suite, puis, parmi les unités 2 et 3, l'unité 3, apparaît comme un îlot qui correspond à la clôture de la construction incrémentale ; elle est réalisée avec une prononciation de type « plateau » (contour intonatif plat) et basse.

La construction fonctionne sur la base d'un suspense à différents niveaux. La première TCU composée d'un syntagme nominal *une interro* suivi d'un qualificatif laisse attendre que l'interro a été l'occasion d'une expérience difficile. Les deux TCU suivantes sont liées par une progression de spécification, à partir d'un élément qui en projette un autre. Une forme de suspense est également maintenue sur le plan prosodique par les deux premières TCU, avant la chute (unité 3), qui formule un constat de désappointement et de déception. Cette construction avec montée d'un suspense puis chute est renforcée par la concision de la dernière unité (*sept*).

Bilan

Dans ces deux analyses succinctes, on retrouve les principes et les arrière-plans fondamentaux des deux approches : le fondement hiérarchique de la macrosyntaxe dans lequel est retraduite la temporalité des constructions ; le caractère éminemment temporel (séquentiel) de la linguistique interactionnelle, ainsi que son caractère « holistique », la syntaxe s'intégrant avec toutes les autres ressources que les participants font jouer, et au même titre qu'elles, dans les constructions.

Deux observations majeures peuvent être émises à partir de ces deux analyses :

- ce que l'on peut considérer comme le point culminant de l'émotion n'apparaît pas au même endroit pour les deux analyses : pour la macrosyntaxe, il se situe entre *une interro qui était censée être facile* et *la moyenne de la classe*, alors que, pour la linguistique interactionnelle, il s'inscrit dans la mise en place progressive d'un suspense qui atteint son sommet avec *la moyenne de la classe*, avant la chute *sept* ;
- l'approche interactionnelle intègre dans l'analyse le contexte dans lequel apparaît la construction (c'est-à-dire le début de l'extrait), en considérant que sa production aussi bien que son interprétation par les participants sont « situées ».

5.1.3 Extrait 2

15	J	• nan mais c'est/
16		(0.2)
17	J	• tous les autres ils étaient blancs:::l ou
18		t' vois <((accélééré)) des couleurs un peu soft/\> • (0.4)
19	->	>ET nous<() ↑WA:::## .h
20	M	• oow ça fait culture quoir•

Analyse macrosyntaxique

Dans l'approche macrosyntaxique, *et nous wah* constitue une autre construction binaire au sens strict, où le prénoyau consiste en un marqueur (appelé « ligateur discursif » par Morel & Danon-Boileau, 1998), *et*, et un pronom, *nous*, et où le noyau n'est pas une construction verbale mais une interjection, *wah*. Il est d'ailleurs parfois délicat, quand le noyau n'est pas une construction verbale, de faire la distinction entre « structure équationnelle » et construction binaire au sens strict. Néanmoins, on notera ici l'impossibilité de postposer le prénoyau (sans le marqueur) : **Et wah nous*. Peut également être remarqué le faux parallélisme, la fausse symétrie, avec les segments précédents *Tous les autres ils étaient blancs ou t' vois des couleurs un peu soft car*, si *tous les autres* est un prénoyau, il ne participe pas d'une construction binaire au sens strict. Ainsi, ce segment est susceptible de devenir postnoyau : *ils étaient blancs ou t' vois des couleurs un peu soft tous les autres*. D'ailleurs, il s'inscrit dans un détachement avec rappel, dans la mesure où il est repris anaphoriquement par *ils*.

Il est intéressant de rapprocher *et nous wah* d'« un énoncé comme *des haricots, bof* [qui] n'a pas d'autre organisation que celle que lui confèrent les positions, les intonations, les ruptures qui en accompagnent la réalisation » (Deulofeu, 1977 : 32).

Analyse interactionnelle

Dans l'approche interactionnelle, la construction binaire prend place dans le cadre d'une autre construction bipartite contrastive *les autres ils étaient...*, *et nous*. Les projections sur lesquelles se fondent ces constructions fonctionnent successivement :

- *les autres ils étaient...* [soit A1] projette quelque chose comme « et le nôtre il était... » [soit A2] ;
- dans le second membre, la première partie *et nous* [soit B1] projette la deuxième partie [B2], qui viendra nécessairement en contraste avec le deuxième membre de *tous les autres ils étaient...*

La construction est du type A [A1=>A2] => B [B1=>B2], où B2 est projeté deux fois : une fois par A2 (par rapport auquel il contraste) et une fois par B1 (projection de type syntaxique). Le contraste entre B2 et A2 est élaboré par différentes ressources (pas seulement sur le plan sémantique, comme serait *blanc vs flashy* ou *vif*) :

- au niveau du lexique, par l'opposition entre l'adjectif (« blanc » suivi d'une expansion) et l'interjection (« WA::: ») ;
- au niveau de la syntaxe, par l'opposition entre, d'un côté, une construction verbale avec un détachement à gauche (« tous les autres ils étaient blancs... »), suivie d'une expansion bien marquée par une accélération et un marqueur phatique (« ou t' vois <((accélééré)) des couleurs un peu soft/> »), et de l'autre, une construction sans verbe (pronom + interjection « >ET nous< (.) ↑WA:::##.h.h ») ;
- au niveau de la prosodie et du rythme, par l'opposition entre un premier membre prononcé lentement et allongé et un second membre très bref - « >ET nous< » est prononcé de façon sèche, comme suspendue (durée de 0.4 sec.) ; l'interjection « wa::: » en revanche, est allongée (0.8 sec.) et prolongée par un souffle. Une pause assez importante sépare les parties (0.4 sec), renforçant encore le contraste.

La construction binaire proprement dite crée moins un effet de suspense que la précédente (il n'y a pas d'effet de chute), mais elle fonctionne également avec un élément fort et concis comme deuxième membre. Ce deuxième membre, commençant plus haut, avec son intonation

montante, son intensité accrue, l'allongement et la voix craquée, a une forte saillance. Il est suivi d'un rire.

Bilan

Les mêmes remarques que précédemment peuvent être émises pour l'analyse de cet extrait. On retrouve l'approche plus locale, et quelque part peut-être plus immanente, de la macrosyntaxe face à l'approche plus située, plus temporalisée et plus holistique de l'analyse interactionnelle.

5.1.4 Extrait 3

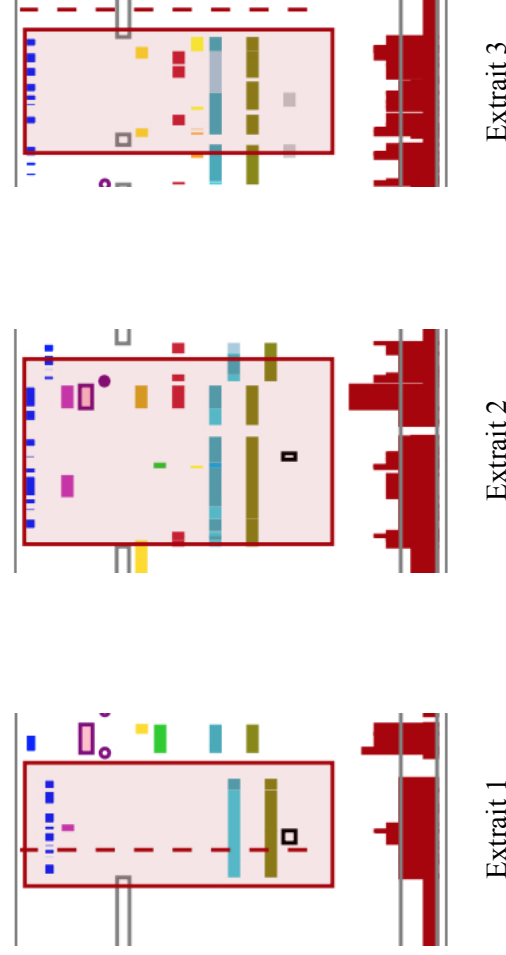
Nous citons ci-dessous un dernier exemple illustrant la présence de marqueurs discursifs dans la construction binaire :

22	J	• tsk et les autres comment ils ont fait • t` sais thierry john et
23		camille (0.5) ils ont fait un truc euh: • ils ont une photo et ça rend
24		bien • mais <((plus vite)) le seul truc c'est que la photo elle a des
25		gros pixels quoi> •
26	M	hm ouais •
27		(0.9)
28		donc ils se foutaient de la gueule de mon jaune/ •
29		(0.3)
30 ->	J	<((en riant)) non mais attends • (.) tes gros pixels là •
31 ->		euh: hein • [tais toi] •
32	M	[[((rire))]]
33		(3.7)

Dans cet extrait, la construction binaire au sens strict est insérée dans du discours rapporté direct. Le premier segment *attends tes gros pixels là euh hein*, borné à gauche de manière externe par le discours citant, est borné de manière interne à gauche par le marqueur *attends* (voir Morel & Danon-Boileau, 1998 ; Groupe ICOR, 2007), et à droite par les marqueurs *là*, *euh*, *hein*. Notons que la présence de *là* et *hein*, vus comme des « ponctuants » de « rhème » par Morel & Danon-Boileau (1998), peut en outre éventuellement inviter à analyser le premier segment comme étant à l'origine un noyau, ensuite recatégorisé en prénoyau devant le noyau *tais-toi*.

5.1.5 Visualisation des indices dans les extraits analysés

La visualisation graphique de ces différents extraits met clairement en évidence un contexte d'accumulation au-dessus de la moyenne, avec des pics, des plateaux et des creux. On peut observer que ce ne sont pas toujours les mêmes indices qui contribuent à ces pics. Dans les extraits 1 et 2, l'accumulation d'indices est portée par la multimodalité/prosodie (allongement ou saillance) ; dans l'extrait 3, ce sont les marqueurs qui compensent d'une certaine façon l'absence d'indices multimodaux. La saillance se positionne dans l'extrait 1 sur le mot *facile* (cf. l'analyse macrosyntaxique) et dans l'extrait 2 sur le « wa:h » (cf. l'analyse interactionnelle).



5.2 Analyse d'un extrait du corpus TCOF

Dans le cadre de l'étude présentée ici, nous avons tout d'abord sélectionné un passage de la base CLAPI afin de tester notre méthode instrumentée pour l'analyse des tonalités émotionnelles dans l'interaction. Nous avons procédé à son analyse et obtenu les résultats mentionnés précédemment. Ensuite, nous avons souhaité mettre à l'épreuve cette méthode sur un autre corpus pour tester sa reproductibilité. Aussi, nous avons choisi un extrait du corpus TCOF : le corpus « Automobile »¹⁴.

Ce corpus a été choisi de la même façon que le corpus présenté ci-dessus, c'est-à-dire en raison de sa potentielle dimension émotionnée. En effet, dans le corpus « Automobile », une locutrice fait le récit à une amie d'un incident survenu lors d'un trajet en voiture. La première raconte à la seconde comment des automobilistes se sont querellés sous ses yeux et le retard que pouvait engendrer cette altercation pour la narratrice qui devait se rendre à des examens à l'université. Ce corpus a également été choisi pour son interactivité, l'émotion ressentie par l'une des locutrices est partagée et coconstruite avec l'autre locutrice au fur et à mesure de l'interaction. La locutrice qui relate les événements alterne entre, d'une part, récit et mise en scène langagière de son récit et, d'autre part, discours rapporté, notamment des insultes proférées par les automobilistes impliqués dans la querelle. Elle explique également qu'au fur et à mesure des événements, elle était de plus en plus stressée, en produisant des énoncés tels que *le stress des partiels qui MONte*.

Le corpus a été annoté dans Praat selon les dimensions décrites précédemment, puis traité par notre outil. En analysant des extraits de ce corpus, nous avons identifié les différentes dimensions susceptibles de participer à la construction d'une émotion. Par exemple, l'extrait suivant comporte des marques multimodales (saillance, débit accéléré), des éléments qui relèvent de l'interactivité (pauses, chevauchements de parole) et des marques d'accord (*tu vois, je vois, ouais, tu sais*) :

L1: tu vois et juste avant [euh]
 L2: [au stop *T1*] ouais
 L1: au stop avant euh soit tu vas vers l'autoroute
 L2: [ouais je vois]

L1: [soit tu vas vers] euh *T1* (0.2) et euh avant juste avant dans le village il y avait euh (0.2) il y avait une caisse devant moi et une autre derrière et celle de derrière lui COLLaît au cul genre VAS-y avance et tout

L2: [ouais]

L1: [mais] super nerveusement quoi et puis on arrive juste avant le stop hop elle double elle se rabat et puis euh ben quand on arrive au stop le matin tu sais il y a une queue

L2: ouais

L'analyse de cet extrait révèle la coconstruction d'une séquence « émotionnée » renforcée par le lexique utilisé par L1. Cette dernière met en scène son récit avec des dimensions langagières et interactionnelles marquant le partage de l'émotion et préparant son interlocutrice au point culminant de son émotion personnelle, le stress d'arriver en retard aux examens.

L'extrait qui suit comporte les mêmes indices, ainsi que des répétitions (*tu vois... tu vois, qui avance qui avance*) :

L1: mais laisse tomber (0.2) elle s'arrête et euh là le mec je je commence il, sort de la caisse et il va voir celui qui l'a doublé et tu sais genre TROP VENERE il commence à faire des gestes dans tous les sens euh tu vois qu'il crie tu vois qu'il est pas content tu sais pas pourquoi mais euh (0.2) voilà et la file de voitures qui avance qui avance et tout d'un coup tu vois il faut il faut quand même qu'il avance donc je fais des appels de phare je klaxonne un petit coup (rires) mais gentil tu vois (rires) hop

L2: tu te dis il a l'air énervé alors je vais peut-être pas trop forcer

L1: (rire) voilà (0.2) et puis euh et puis le stress des partiels qui monte oh là là il faut quand même que [ça

L2: [je vais] être en retard

L1: ça avance et là je vois le type qui retourne pour aller dans sa caisse donc je me dis bon c'est cool voilà (0.2) ça c'est fait et euh à mon avis le mec qui était dans la caisse de derrière a dû faire ouais c'est ça connard tu vois

L2 : ouais

L'analyse de cet extrait montre également que nous pouvons identifier ici une séquence « émotionnée ». Il semble que la combinaison des dimensions décrites en 3 assure la présence d'une (ou plusieurs) émotion(s) dans le discours des locutrices. En effet, chaque dimension prise isolément n'est pas un marqueur d'émotion mais c'est bien la présence cumulée de différentes dimensions qui participe à la coconstruction d'une tonalité émotionnelle.

Fort de ces analyses, nous avons comparé les séquences ainsi identifiées à la courbe émotionnelle proposée par l'outil. Comme pour l'extrait « Jaune Fluo », nous constatons que les pics repérables grâce à la représentation graphique correspondent bien aux séquences à tonalité émotionnelle dégagées par l'analyse. L'outil nous permet de prendre du recul par rapport au lexique afin de repérer des passages présentant potentiellement une telle tonalité. La méthode mise en œuvre, et exposée au début de cet article avec le passage de la base CLAPI, nous semble donc bien reproductible à d'autres corpus.

6 Conclusion et perspectives

Nous avons proposé une méthode et conçu un outil exploratoire pour une approche interactionnelle de l'émotion dans les interactions. Cette méthode repose sur le repérage d'indices d'émotionnalité à travers huit dimensions d'analyse : la multimodalité, l'interactivité, les bruits-sons, les marques d'accord et de désaccord, les marques de personnes, les unités macrosyntaxiques, les unités interactionnelles et les répétitions. Deux extraits de corpus provenant de deux équipes différentes ont été analysés suivant cette approche. Nous avons

constaté que l'outil de visualisation permet d'identifier et de mesurer les phénomènes de concomitance et d'accumulation d'indices à un instant donné, mais aussi de faciliter l'analyse en isolant chaque indice pertinent.

Concernant l'appréhension de l'émotion dans l'interaction, il conviendra, par la suite, d'améliorer le modèle proposé afin qu'il puisse rendre compte des effets de séquence ou d'accumulation dans le temps (phénomènes d'inertie ou de contraste) et permette également de pondérer les niveaux d'annotations. Par ailleurs, pour des raisons essentiellement méthodologiques, la contribution du lexique a été volontairement exclue des dimensions d'analyse : une approche complémentaire consisterait à intégrer le calcul de l'impact de chaque mot dans une approche lexicographique (voir Plantin, 2015b).

Références bibliographiques

- Astésano, C., Espesser, R. & Rossi-Gensane, N. (2008). Quelques cas particuliers de détachement à gauche – ou la prosodie à l'atde de la syntaxe. *Actes des XXVII^{èmes} Journées d'Étude sur la Parole*. Juin 2008, Avignon, France.
- Blanche-Benveniste, C., Bilger, M., Rouget, C. & Van den Eynde, K. (1990). Le français parlé. Études grammaticales. Paris : CNRS Éditions.
- Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 6.0.08, retrieved 5 December 2015 from <http://www.praat.org/>
- Bruxelles, S. & Traverso, V. (2001). La particule « ben » dans deux type d'interactions. In Pons Borderia, S. (éd.), *Numéro spécial de la revue Cuadernos de Filología*.
- Bruxelles, S. & Traverso, V. (2006). Usages de la particule « voilà » dans une réunion de travail : analyse multimodale. In Drescher, M. & Frank-Job, B. (éds), *Les marqueurs discursifs dans les langues romanes : approches théoriques et méthodologiques*. Francfort : Peter Lang, 71-93.
- Couper-Kuhlen E. (1996). The prosody of repetition: On quoting and mimicry. In Couper-Kuhlen, E. & Selting, M. (éds), *Prosody in conversation*. Cambridge : Cambridge University Press. 366-405.
- Denis, A., Quignard, M., Fréard, D., Détienne, F., Baker, M. & Barcellini, F. (2012). Détection de conflits dans les communautés épistémiques en ligne. *TALN - Actes de la Conférence sur le Traitement Automatique des Langues Naturelles - 2012*. Juin 2012, Grenoble, France. <hal-00766397>
- Deulofeu, J. (1977). La syntaxe et les constructions binaires. *Recherches sur le français parlé*, 1, 30-61.
- Ekman, P. (1999). Basic emotions. In Dalgleish, T. & Power, T. (éds), *The handbook of cognition and emotion*. Chichester, UK : John Wiley, 45-60.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86-88. doi: 10.1126/science.164.3875.86
- Groupe ICOR (Balthasar, L., Bruxelles, S., Mondada, L. & Traverso, T.). (2007). Variations interactionnelles et changement catégoriel : l'exemple de « attends ». In Auzanneau, M. (éd.), *La mise en oeuvre des langues dans l'interaction*. Paris : L'Harmattan, 299-319.
- Groupe ICOR (Bert, M., Bruxelles, S., Étienne, C., Mondada, L. & Traverso, V.). (2008). Tool-assisted analysis of interactional corpora: 'voilà' in the CLAPI database. *Journal of French Language Studies*, Vol. 18/01, 121-145.
- Groupe ICOR (Étienne, C., Bruxelles, S., Jouin, É., Mondada, L., Oloff, F., Teston, S. & Traverso, V.). (à paraître). Les hétérorépétitions dans une interaction orale : définition et conception d'un outil de détection automatique. In Richard, É. & Oriez, S. (éds), *Des organisations dynamiques de la langue orale*. Berne : Peter Lang.
- Hamon, T., Fraise, A., Paroubek, P., Zweigenbaum, P. & Grouin, C. (2015). Analyse des émotions, sentiments et opinions exprimés dans les tweets : présentation et résultats de l'édition 2015 du défi fouille de texte (DEFT). *Actes de la 11e Défi Fouille de Texte DEFT 2015*. Caen : ATALA.
- Heinemann, T. & Traverso, V. (éds) (2009). Complaining in interaction. *Journal of Pragmatics*, 41(12), 2381-2578. doi: 10.1016/j.pragma.2008.10.006

- Hepburn, A. & Potter, J. (2012). Crying and crying responses. In Peräkylä, A. & Sorjonen, M.-L. (éds), *Emotion in Interaction*. Oxford : Oxford University Press, 195-211.
- Hoey, E. (2014). Sighing in interaction : somatic, semiotic, and social. *Research on Language and Social Interaction*, 47, 2, 175-200.
- Holt, E. (2011). On the nature of “laughables” : laughter as a response to overdone figurative phrases. *Pragmatics*, 21,3, 393-410.
- Kerbrat-Orecchioni, C. (2000). Quelle place pour les émotions dans la linguistique du XXe siècle ? In Plantin, C., Doury, M. & Traverso, V. (éds), *Les émotions dans les interactions*. Lyon : Presses Universitaires de Lyon, 33-74.
- Kerbrat-Orecchioni, C. (2005). *Le discours en interaction*. Paris : Armand Colin, 50-55.
- Kerbrat-Orecchioni, C. (2012). Analyser du discours : le cas des débats politiques télévisés. *3^{ème} Congrès Mondial de Linguistique Française – CMLF 2012*, 25-42. DOI 10.1051/shsconf/20120100338
- Lacheret, A., Kahane, K., Pietrandrea, P., Avanzi, M., & Victorri, B. (2011). Oui mais elle est où la coupure là ? Quand syntaxe et prosodie s'entraident ou se complètent. *Langue française*, 170, 61-80.
- Martin, P. (1996). WinPitch: un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows. *Actes des XXIV Journées d'Étude sur la Parole, Avignon, mai 1996*. 224-227.
- Morel, M.-A. & Danon-Boileau L. (1998). *Grammaire de l'intonation. L'exemple du français oral*. Paris : Ophrys.
- Peräkylä, A., & Sorjonen, M.-L. (éds) (2012). *Emotion in interaction*. Oxford : Oxford University Press.
- Plantin, C. (2011). *Les bonnes raisons des émotions*. Berne : Peter Lang.
- Plantin, C. (2012). Les séquences discursives émotionnées : Définition et application à des données tirées de la base CLAPL. *3^{ème} Congrès Mondial de Linguistique Française. Section Discours, Pragmatique et Interaction*. <http://dx.doi.org/10.1051/shsconf/20120100218>
- Plantin, C. (2015a). Emotion and affect. *The International Encyclopedia of Language and Social Interaction. First Edition*. Karen Tracy (General Editor), Cornelia Ilie & Todd Sandel (Associate Editors). Chichester : John Wiley. DOI: 10.1002/9781118611463/wbielsi079
- Plantin, C. (2015b). Paura, emozione, passione, sentimento : Étude de la contagion émotionnelle d'après le Dizionario Combinatorio Itahned. *Le langage et l'homme*, 50, 2, 43-58.
- Plantin, C., Doury, M. & Traverso, V. (2000). *Les émotions dans les interactions*. Lyon : Presses Universitaires de Lyon.
- Prévost, S. (2003). Détachement et topicalisation : des niveaux d'analyse différents. In Neveu, F. (éd.), *Linguistique du détachement, Cahiers de Praxématique*, 40, 97-126.
- Rastier, F. (1987). *Sémantique interprétative*. Paris : Presses Universitaires de France.
- Schegloff, E. (2007). *Sequence Organization in Interaction. A primer in Conversation Analysis*. Cambridge : Cambridge University Press.
- Tolins, J. (2013). Assessment and Direction through Nonlexical Vocalizations in Music Instruction. *Research on Language and Social Interaction*, 46, 1, 47-64.
- Traverso, V. (2005). Quelques formats intégrant la répétition comme ressource pour le développement thématique dans la conversation ordinaire. In Galatolo, R. & Fasulo, A. (éds), Special Issue on Conversation Analysis. *Rivista di Psicolinguistica Applicata*, 153-166.
- Wundt, W. (1911). *Völkerpsychologie. Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte* (Vol. 1 : Die Sprache). Leipzig : Verlag Wilhelm Engelmann (réédité en 1975 chez Scientia Verlag à Aalen).

¹ Nous tenons à remercier les relecteurs de leurs commentaires. Ce travail a été réalisé dans le cadre du projet ANR ORFÉO, Outils et Recherches sur le Français Écrit et Oral (<http://www.projet-orfeo.fr>) coordonné par J.-M. Debaetsueu, au sein d'une étude-pilote sur l'émotion. Si le but général du projet ORFÉO est de rassembler des outils interopérables d'annotation de corpus et des ressources orales et écrites dans des formats homogènes, ORFÉO offre

aussi un cadre pour mener des recherches innovantes sur les corpus, rendues possibles par la mise en commun de ces ressources et outils, et pour envisager des développements spécifiques pour des besoins nouveaux. L'étude-pilote « Émotion », pour laquelle cet article fait état des premiers résultats, a donc tiré parti à la fois du cadre collaboratif de ce projet (l'étude-pilote a impliqué deux équipes pluridisciplinaires) et des ressources/outils développés par d'autres partenaires du projet.

² <https://defl.limsi.fr/2015/>

³ Base CLAPI : <http://clapi.ish-lyon.cnrs.fr>

⁴ Base TCOF : <http://www.cnrtl.fr/corpus/tcof/>

⁵ Base CLAPI : <http://clapi.ish-lyon.cnrs.fr> – Corpus : *Repas ~ conversations entre étudiants - lyon 2006*. Enregistrement : *Repas français: vraiment vraiment désolée*. Le corpus figure dans les ressources du projet ORFÉO et sera donc disponible également sur Ortolang (<https://www.ortolang.fr>) à la fin du projet (2016).

⁶ L'annotation des phénomènes prosodiques a été réalisée selon l'approche interactionnelle, de manière perceptuelle. Il n'a donc pas été fait appel à des outils d'analyse tels que Praat ou WinPitch (Martin, 1996). Par la suite, d'autres paramètres (comme les frontières intonatives, par exemple) pourront également être pris en compte, perceptuellement ou automatiquement.

⁷ La reconnaissance de locuteurs (*diarisation*) est une tâche dont le but est de reconnaître qui parle et quand. Elle se différencie en cela de la reconnaissance de la parole qui cherche à identifier ce qui est dit.

⁸ Bruxelles & Traverso, 2001 ; Bruxelles & Traverso, 2006 ; Kerbrat-Orecchioni, 2005 ; Groupe ICOR, 2007.

⁹ Dans le projet ORFÉO, les interjections et particules sont réparties dans plusieurs catégories (interjections, conjonctions de coordination/subordination, adverbes, prépositions...) alors que les marqueurs comme *tu vois* ou *je veux dire* sont traités comme *tu vois*, c'est-à-dire clitique + verbe par l'analyseur microsyntactique.

¹⁰ La notion de rection est entendue au sens de Blanche-Benveniste *et al.* (1990 : 291) comme « l'ensemble des éléments [...] organisés syntaxiquement par le verbe ».

¹¹ Une exception à cette règle : les marques de frontières (enchaînement immédiat, alternance de locuteurs). Ces marques, qui n'ont pas de durée chronométrée, se voient attribuer une durée arbitraire de 280ms. Cette valeur correspond graphiquement à celle d'un cercle de 7 pixels de rayon.

¹² JTrans est un logiciel d'alignement de transcription. Disposant de toute la transcription, il exploite des techniques de reconnaissance de la parole pour détecter la position des phonèmes dans le temps. Il est librement téléchargeable : <https://gforge.inria.fr/projects/jtrans/>. Précisons que cet outil est utilisé dans le projet ORFÉO.

¹³ http://icar.univ-lyon2.fr/projets/corinte/documents/2013_Conv_ICOR_250313.pdf

¹⁴ Le corpus Automobile est accessible par le lien suivant : <http://www.cnrtl.fr/corpus/tcof/>. Ce corpus est recensé dans le corpus ORFÉO.