

LIMSI-CNRS - Traitement du Langage Parlé

Thème: « Dimensions affectives et sociales des interactions parlées »



Master 2 Recherche PXI

Les émotions dans la communication:
analyse et détection

Laurence Devillers
Professeur Paris-Sorbonne PIV
devil@limsi.fr







Team: « Affective and social dimensions of spoken interactions »

L. Devillers – Professor at Sorbonne University – LIMSI-CNRS(devil@limsi.fr)
C. Chastagnol, A. Delaborde, M. Sehili, M. Soury, M. Tahon, F. Yang
http://www.limsi.fr/tlp/topic2.html

Since 2001 team 5-10 researchers (4 PHD students)

- Detection and analysis of affective and social components of audio signals exchanged between a human and his partner (human, robot or virtual agents)
- Dynamic speaker profil modeling (affective and interactionnel dimensions such as extraversion...) to adapt the strategy of a dialog system
- Combinaison of verbal and non verbal audio and audiovisual information
- Real-life emotions mixtures of emotion (pos/neg) in H-H dialogs
- Induction techniques (WoZ) with machine (robots or agents)

Cadre du cours

Ce cours s'attache à présenter les théories relatives aux émotions selon les perspectives psychologiques et neurologiques et enfin à présenter comment l'informatique :

- aide à la connaissance :
 - comprendre les comportements humains, affectifs et sociaux
- construire de nouvelles interactions :

interfaces Humain-Machine

applications: remédiation, compagnon artificiel, jeux, aide au diagnostic

Davillam Master 2 2014

Plan

- 1. Introduction
- 2. Projets au LIMSI Sujets de stage
- 3. Théorie des émotions
- 4. Corpus
- Annotation
- 6. Indices émotionnels dans la parole
- 7. Détection des émotions dans la voix

L. Devillers Master 2 2014

L. Devillers Master 2 2014

« Qu'est ce qu'une émotion ? »

- « Everyone knows what an emotion is, until asked to give a definition.
 Then it seems, no one knows » [Fehr et Russell 1984]
- «... phénomène empirique, généralement transitoire et d'une certaine intensité qui se manifeste par des indices dans la voix, le visage, les gestes etc. et par des changements physiologiques » [Caffi et Janney 1994]
 - → 'émotion' pris au sens large « état affectif »

Emotion : codage à différents niveaux

Partie 1:

Introduction

Perception: Emotion perçue

Emotion exprimée

Cognition: Emotion interne cachée

L. Devillers Master 2 2014

L. Devillers Master 2 2014

Émotions dans la communication

Les émotions sont « les interfaces de l'organisme avec le monde extérieur » [Sherer 81].

3 fonctions principales:

- réaction individuelle à un stimuli particulier,
- préparation physiologique et psychologique de l'organisme à une action appropriée,
- fonction de communication de ses intentions à un autre organisme

L. Devillers Master 2 2014

Intelligence Emotionnelle

- la connaissance de soi : le fait de pouvoir identifier ses émotions
- la maîtrise de ses émotions : la capacité d'adapter ses sentiments à chaque situation,
- l'auto-motivation : le fait d'être capable de remettre à plus tard la satisfaction de ses désirs et de réprimer ses pulsions,
- · la perception des émotions d'autrui : l'empathie,
- la maîtrise des relations humaines: savoir entretenir des relations courtoises avec les autres.

Perspectives théoriques

Origine de la parole : Transformations anatomiques. Les premiers mots $% \left(1\right) =\left(1\right) \left(1$

seraient des onomatopées.

Origine de la musique : Cohésion sociale

Loi de Spencer (1911) : une émotion ressentie influence les caractéristiques de l'expression verbale et musicale du sujet.

Cette assertion constitue le point de départ de l'analyse des émotions en musique et en parole

Conceptualisation des émotions

Il existe un nombre réduit d'émotions desquels découlent toutes les autres émotions : palette émotionnelle

Les basic emotions : joie, tristesse, peur, dégoût

L. Devillers Master 2 2014

QI + IE: Intelligence émotionnelle: Salovey & Mayer, 1990

Deux sortes d'intelligence:

- intellectuelle (néocortex régions supérieures du cerveau) et
- émotionnelle (région inférieures et plus anciennes du cerveau dans le sous-cortex)

IE: centres émotionnelles + coordination avec les centres intellectuels

L. Devillers Master 2 2014

L. Devillers Master 2 2014

Les systèmes Homme-Machine "Intelligent Affectivement"

Puisque les émotions jouent un rôle central dans les interactions sociales

-> les machines de demain devront avoir une intelligence émotionnelle et sociale

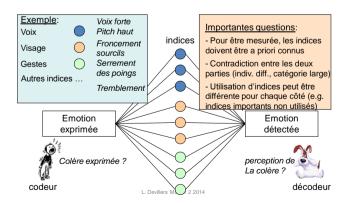
"Affective computing" (1997), laboratoire MIT MEDIA lab (R. Picard)

Sciences affectives (2005), UNIGE (K. Scherer)

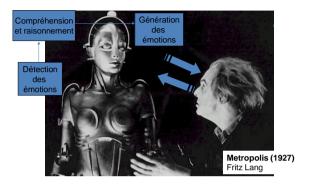
Réseau d'excellence HUMAINE (2004 - 2008), coordinateur (R. Cowie) CNRS LIMSI, Cergy, Univ. Paris 8 et FT, Cantoche

L. Devillers Master 2 2014

Communication des émotions



Affective computing



L. Devillers Master 2 2014

Modélisation informatique

But : Modèle quantitatif des émotions

Analyse objective du comportement émotionnel par l'expression de l'émotion dans la voix, le geste, le visage...

Détection des émotions

- Détecter des situations anormales pour adapter les stratégies des systèmes de dialogue H-M,
- Améliorer la reconnaissance de parole spontanée,
- · Piloter un agent à la voix (second-life),

Génération des émotions

ACAs: agents conversationnels animés

Traitement automatique des émotions

Plus avancé en génération qu'en détection

Agents conversationnels

Modèle OCC, Ortony, Clore & Collins (1988) :

Schéma de classification des émotions (26) basé sur une réaction de valence aux évènements et objets selon les buts, les standards et les attitudes d'un individu.

- Animation FACS norme ISO MPEG 4
- APML: affective presentation markup language (Greta PVIII)
- Réalisations: Projets d'animation (NECA, MAUI, ...)
- Robotique (Chien AIBO Sony)

Réseau d'Excellence FP6 HUMAINE/AAAC (2004-2008-) :

Terminologie commune, Corpus de travail commun (multimodal) Standard d'annotation, d'évaluation Modèles informatiques permettent de vérifier des théories cognitives

L. Devillers Master 2 2014

Pourquoi détecter les émotions?

Objectif global : analyser et détecter/prédire les

comportements émotionnels

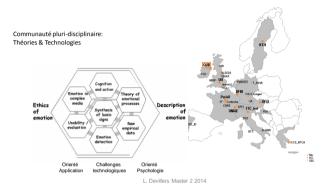
- les émotions peuvent être détectées dans le visage (Ekman : émotions de base sur des images), la posture, la voix...
- émotions dans la voix étudiées au Limsi depuis 2001 (données d'interaction (REAL-LIFE) riches en émotion)

Applications

- CHM: centres d'appels, ACA, robotique
- Fouille de données: indexation par le contenu émotionnel
- Mesure de l'émotion: médical, marketing
- Sécurité: audio-vidéo surveillance, filtre

Réseau d'excellence HUMAINE

emotion-research.net



Données réelles en contexte Centre d'appels

"hot line": assistance technique TV, cable et internet (US)









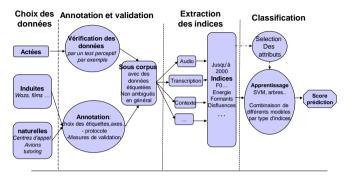






L. Devillers Master 2 2014

Méthodologie: Système de détection des émotions



L. Devillers Master 2 2014

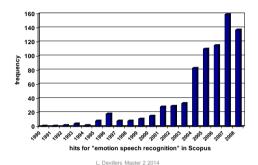
Quelques applications

Futures interfaces évoluées orientées émotions

- Multiples domaines: marketing/publicité/commerce, design émotionnel, jeux, films, réalité virtuelle, etc. centre d'appels, robotique
- Objectifs stratégiques :
 - · Sécurité (transports: route, air, audio-vidéo surveillance)
 - · Santé (bio-senseurs, professions médicales)
 - · Apprentissage (pédagogie)
- Réalisations :
 - · Animation de scènes/robotique
 - · Capteurs sensoriels vêtements (MIT)

L. Devillers Master 2 2014

Emotion et Parole Ce n'est plus un sujet exotique ! (Batliner 2009)



Exemples d'applications



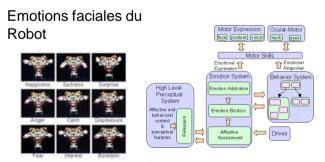
Serious Game Modification du comportement

Raisonnement:

- Buts et croyances, Ex: [Marsella2002]
- Attention, Ex: [Lang and Davis, 2006]
- Mémoire, Ex: [Velasquez1998]

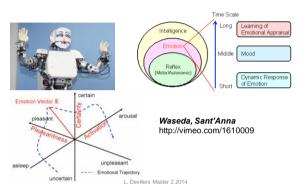


Emotion et robot (KISMET MIT 2007)



KISMET (MIT)
http://www.youtube.com/watch?v=dKZczUDGp_I
L. Devillers Master 2 2014

Emotion et robot (Waseda Face Robot - ROBORAMA2005)



Emotion et robot (Leonardo MIT 2007)

Leonardo: The most expressive robot in the world today!



robots with social intelligence that communicate and learn the same way people do

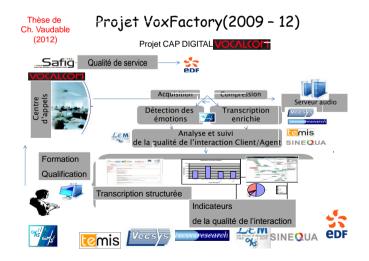
L. Barsalou, C. Breazeal & L. Smith (2007) "Cognition as Coordinated non-cognition." Cognitive Processing 8: 79-91.

L. Smith and C. Breazeal (2007) "The Dynamic Life of Developmental Process." Developmental Science, 10(1), 61-68.

http://www.youtube.com/watch?v=xYSmp3bjP_0&feature=related

Partie 2 Projets au LIMSI - TLP

L. Devillers Master 2 2014



Projet Affective Avatar (2008-10)

automatique Détection





Projet ANR Affective Avatar



Mais tu fais quoi toi pendant ce temps (énervement) Ben ouais tu fais quoi (énervement) Tu comptes les points c'est facile (colère)

A chaque fois qu'il y a un problème tu laisses tomber et Puis après tu critiques (colère)

T'as qu'à le changer le monde toi puisque tu es si pur (colère



Annotations Animations Colère faible

Thèse de M. Tahon (2012)

Projet ROMEO (2009 - 12)

A. Delaborde Projet CAP DIGITAL ROMEO http://www.projetromeo.com B. (2013) ALDEBARAN



NAO 58cm -> ROMEO 1m40

Le projet vise à développer un robot humanoïde destiné à devenir un véritable assistant des personnes en perte d'autonomie.

Etude également du robot comme compagnon de jeu : première étude avec NAO.



acapela



angel





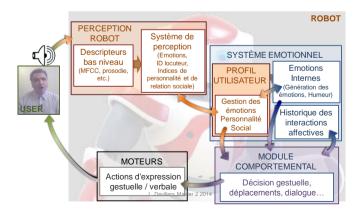
LIMSI : - Id du locuteur et de ses émotions

- Système émotionnel

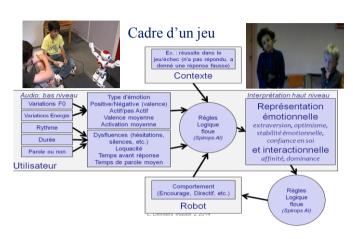


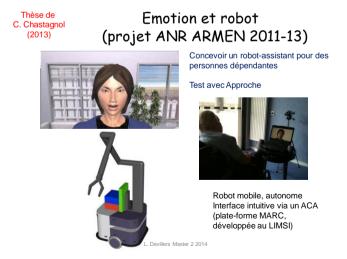


Emotion et robot (projet FUI ROMEO 2010)



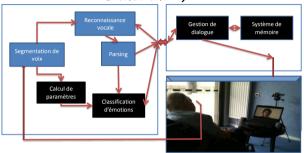
Interaction sociale Homme-Robot (LIMSI)





Dialogue intuitif

avec l'interface homme-robot représenté par l'agent MARIE - parole spontanée (Agent LIMSI Marc)



L. Devillers Master 2 2014

Thèse de M. Soury (en cours)

Projet FEDER E-Therapies (oct. 2011-14)

□ Partenaires : centre-émotion - Pitié-Salpêtrière (A. Pelissolo, R. Jouvent)

LIMSI – tlp : détection des émotions dans l'audio

PME (Voxler, Tekneo)

☐ Outils pour soigner les phobies sociales (serious game):

- Les thérapies cognitivo-comportementale se basent sur une mise en situation de la personne avec un environnement anxieux de manière graduée
- > Verrous technologiques :
 - la mesure de l'état d'anxiété/stress de la personne (indices audio-visuels);
 - la gestion d'un dialogue oral entre la personne et le « répondant » virtuel (ex : avatar 3D);
 - · Bio-feedback de l'état détecté;
 - · la gestion du caractère d'évitement, mesure de l'attention;
 - le suivi de l'apport thérapeutique (interface avec le thérapeute)



Projet ANR COMPARSE (janv. 2012-15)

- ☐ Etude du stress dans la prise de parole en public (sujets non pathologiques/pathologiques (bi-polaire))
 - > Verrous technologiques :
 - La mesure de l'émotion/stress de la personne (indices audio, visuels, gestuels et physiologiques);
 - · Comparaison situation avec agent/sans agent
 - > Cadre applicatif: entretiens d'embauche (étudiants, patients)
 - · Protocole de collecte dans une situation « écologique »
 - Stress prise de parole, défaut/qualité, situation inattendue, public évaluatif, enjeu de compétition
 - Capture peu invasive ex: audio micro-cravate, kinect, caméra (sujet/public), oxymètre
 - Même test avec public/sans, avec différents comportements d'avatars



Nouveaux projets





Thèse de F. Yang (en cours)







ROMEO 2 : interaction homme-robot pour les personnes dépendantes (2013-17)

CHIST-ERA JOKER: un robot qui fait de l'humour (2014-17)

Sujets de stages

- Deep Neural Networks for acoustic emotion and intention recognition (CHIST-ERA JOKER)
- User-model in H-Robot social interaction with multimodal perception (Projet OSEO ROMEO2)
- Effective long-term interaction by virtue of establishing an empathic relationship between system and user (CHIST-ERA EUPHORIA)

EU Social interactions with robots, New Chist-era (en cours)

Partie 3: Théorie des émotions

Antiquité: manuels de rhétorique (Aristote, Ciceron), émotions dans le langage oral

Pluri-disciplinarité

- Psychologie
- Psychiatrie
- · Neurologie
- · Biologie
- Philosophie
- Linguistique
- · Traitement du signal acoustique
- · Traitement d'image
- · Traitement d'indices physio
- Informatique
- · Intelligence artificielle
- · Reconnaissance des formes

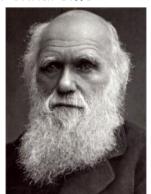
L. Devillers Master 2 2014

Le fonctionnalisme

À la fin du XIX ième siècle la théorie de l'évolution de Darwin commence à avoir une certaine influence en psychologie.

Dans la notion de « survie du plus fort » Darwin insiste sur le rôle que jouent les structures biologiques supérieures dans l'adaptation des organismes à leur environnement.

Réf.: L'expression des émotions chez l'homme et les animaux (1872)



Charles Robert Darwin Naturaliste anglais (1809-1882)

L. Devillers Master 2 2014

Résumé: Approches théoriques des émotions (Cornelius 2000)

Théorie Darwiniste (1872) - Ekman (1978), Plutchick (1970)

- Émotions ont des fonctions de survie (~mêmes pour toutes les espèces)
- > Universalité d'un petit nombre d'émotions primaires : Darwin 5, Ekman 6,... Théorie de James(1884) — Damasio(1994)
- ➤ Interdépendance entre processus cognitifs, physiologiques et émotionnels Théorie cognitive Lazarus(1968), Zajonc (1984), Sherer(1981)
- Modèle d'évaluation cognitive : Organisme évalue les evts sur la base d'une série de dimensions d'évaluations cognitives séquentielles du stimulus (pertinence, implication, potentiel de maîtrise.

Théorie socio-constructive - Averill(1980) - Shweder(1993)

 Contexte socio-culturel et la situation influence la perception des émotions.

L. Devillers Master 2 2014

Darwin et James

Théorie Darwinienne (1872)

Émotion : vestige chez l'être humain des fonctions de survie. Plutchick (1970)

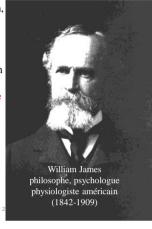
8 émotions primaires (comportements primitifs):
 confiance (incorporation), dégoût (rejet),
 peur (protection), colère (destruction),
 joie (reproduction), tristesse (réintégration),
 surprise (orientation), curiosité (exploration)

Théorie biologique 19ième (James)

Émotion : perception des changements physiologiques, dus aux stimuli L Devillers Master 2 2014

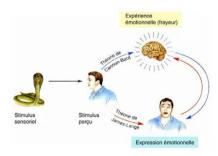
- Inspirée par les idées de Darwin, le fonctionnalisme étudient la <u>fonction</u> des processus psychiques qui permettent à l'individu de s'<u>adapter</u>.
- Grande influence sur l'évolution de la w.
- Recherche sur les émotions et le comportement observable
- Apprentissage, Questionnaires, Tests psychologiques
- Réf. Article: Qu'est-ce que une émotion ? (1884)

L. Devillers Master 2



La théorie centrale des émotions de Cannon

Théorie de Cannon-Bard: l'expérience émotionnelle est indépendante de l'expression émotionnelle. Selon la théorie de Cannon-Bard (flèches bleues), la frayeur résulte de la perception du stimulus, et ensuite seulement il y a une réaction comportementale.



De James à Zajonc

James (1884): S'enfuit-on à cause d'un ours parce qu'on a peur ou a-t-on peur parce qu'on s'enfuit?

Pour lui, les processus viscéraux étaient la cause de l'émotion qui naissait seulement quand le cerveau en prenait conscience.

Shachter - Singer (1960): émotion surgit lorsque nous attribuons grâce à nos capacités cognitives, une explication à des signaux corporels ambigus

Arnold - Lazarus (1980) : notion d'évaluation d'une situation. l'interprétation d'une situation influence fortement l'émotion ressentie. Rapprochement entre émotion et cognition.

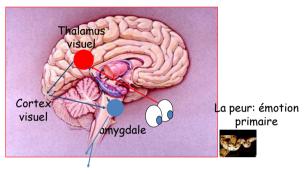
Zajonc (1984) : montre que les émotions peuvent être indépendante de la cognition.

-> Etudes actuelles sur la perception inconsciente des émotions. Nos réactions émotives peuvent survenir en l'absence de la conscience explicite d'un stimulus.

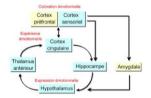
L'émotion des neurosciences Cognitives (1)

- Comment: Etude de patients atteints de maladies neurologiques et de lésions cérébrales et Imagerie cérébrale IRM (résonance magnétique fonctionnelle)
- J. LeDoux, neuroscience
 Réf. Le cerveau émotionnel (1996), neurobiologie de la personnalité (2003)
- A montré le rôle de l'amygdale dans le système de défense ou de peur, elle détermine les actions en cas de menace
- Emotion: état du cerveau particulier associé à une réponse corporelle donnée.
 Niveau d'analyse des émotions, sous-symbolique (inconscient) Sentiments: interprétations conscientes des processus réels que sont les émotions.

La peur



L. Devillers Master 2 2014



Les signaux sensoriels

-> relai distributeur (thalamus) -> ensemble de structures cérébrales entre l'hypothalamus et le néocortex (voie courte)

-> cortex cérébral (voie longue)

Voie courte : voie émotionnelle, permet à l'organisme de préparer l'action avec une réponse stéréotypée (innée ou acquise), rapidité

Voie longue : voie rationnelle, permet une analyse élaborée de la situation pour pondérer les informations

2 analyses interdépendantes, émotion essentielle sans elle pas de mémorisation, ni de prise de décision

L. Devillers Master 2 2014

L'émotion des neurosciences Cognitives (2)

• Principalement expérimentations cliniques. A. Damasio, neurologue américain

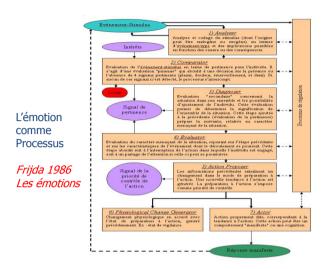
Réf: L'erreur de Descartes (1994), Spinoza avait raison (1999), le sentiment même de soi (2003)

- Il affirme que « la capacité à s'exprimer et à ressentir des émotions est indispensable à la mise en œuvre des comportements rationnels », sans émotion l'homme ne peut vivre en société. Interdépendance des processus cognitifs, physiologiques et émotionnels
- Technologies de l'imagerie cérébrale vont permettre de révéler la continuité du processus qui commence par un stimulus, se poursuit par des réponses émotionnelles et leur représentation et se conclut par la perception consciente que nous appelons sentiments.

L. Devillers Master 2 2014

L'émotion des psychologues

- · Qu'est-ce qu'une émotion ? Une émotion associe :
 - · Des sensations physiques,
 - Des représentations mentales,
 - · Des réponses motrices spécifiques
- L'importance et la place de ces éléments dans la survenue de l'émotion varie suivant les approches.
- Pour Sherer (1987), les émotions ne sont pas des états simples mais une séquence dynamique d'états variables



Théories de l'interaction

ex: Théorie des faces (Goffman)

- · Socio linguistique
- Rites d'interaction: canevas d'actes verbaux et non verbaux:
- Face: valeur sociale positive
 - Sauver la face, perdre la face
- · Figure: faire bonne figure
- Figuration (masque) : comédie sur la scène sociale
- · Régles de bonne conduite:
 - Règles d'amour-propre
 - Règles de considération

ex: Modèle push-pull (Scherer)

L. Devillers Master 2 2014

La théorie d'Averill : un point de vue du constructivisme social

Dans les précédentes théories, les émotions ont été envisagées dans une perspective intrapsychique.

- Pour Averill (1980): les émotions seraient le produit de « constructions sociales » et dépendraient essentiellement du contexte social dans lequel elles apparaissent.
- L'émotion est essentiellement déterminée par les normes, les règles sociales qui existent dans un environnement social donné.
- L'évaluation émotionnelle du stimulus correspond à des jugements hédoniques personnels sur ce qui est désirable ou indésirable.
- L'émotion n'est pas quelque chose que les individus font (comme une action) mais qu'ils subissent (comme une passion). C'est une expérience de passivité.

L. Devillers Master 2 2014

Partie 4: Théorie Annotation Corpus

Un challenge avec des corpus naturels...

What's basic about basic emotions?

Psychological Review, 97, 315-331 - Ortony(1990)

Arnold : colère, aversion, courage, découragement, désir, désespoir, peur, haine, espoir, amour, tristesse

Darwin : colère, peur, joie, tristesse, dégoût

Ekman: colère, peur, joie, tristesse, dégoût, surprise

Frijda: désir, intérêt, bonheur, surprise,

Izard: colère, mépris, dégoût, détresse, peur, culpabilité,

intérêt, joie, honte, surprise

James: peur, douleur/chagrin, amour, rage

Mowrer: souffrance, plaisir,

Oatley: colère, dégoût, inquiétude, bonheur, tristesse

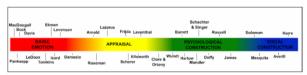
Plutchik: curiosité/surprise, dégoût/confiance, ,joie/tristesse,

peur/colère

L. Devillers Master 2 2014

Théories et modèles des émotions placés sur un continuum (Gross et Barrett 11)

Dans des situations spontanées, il est fréquent d'observer des mélanges d'émotions (Scherer, 98, Devillers 05). Baron-Cohen propose une liste plus globale de 416 « états mentaux » incluant des états comme la relaxation ou l'incertitude (Golan et al., 06).



Des études montrent que les personnes ajustent leurs comportements en se basant sur l'interprétation des émotions des personnesqui les regardent (Parkinson, 10).

L. Devillers Master 2 2014

Etats affectifs (Sherer)

Intensité

Emotions (ex: colère, tristesse, joie, peur, honte, fier, exalté, désespéré)

Humeurs (ex: gai, pessimiste, apathique, irritable, dépressif, optimiste)

Positions entre personnes (ex: distant, froid, méprisant, chaleureux)

Attitudes (ex : amour, haine, amitié, désir)

Dispositions affectives (ex: nerveux, anxieux, morose, hostile)

Durée

- Caractéristiques pour différencier les états affectifs, ex:
- Réponses : durée et intensité
- Conséquences: rapidité de changement (émotions, positions entre personnes, humeurs...), impact comportemental (émotions, attitudes, ...)
- > Etats générés de façon complexes : push/pull effects
 - > biologiques, + différence : idiosyncrasie (réaction innée) /contexte
 - > effet socio-culturel différence : codes

L. Devillers Master 2 2014

Annotation de corpus 3 Théories de représentation des émotions

- Etiquettes émotions :
 - Émotions primordiales instinctives associées au systèmvégétatif: faim, soif, etc. (Denton)
 - 6 émotions de base (Ekman) primaires innées et universelles: peur, colère, joie, tristesse, surprise, dégo
 - Émotions secondaires ou sociales : embarras, jalousie, culpabilité

Plutchik soumission= acceptation + peur Description plus générique fondée sur des dimensions

- Description plus générique fondée sur des dimensions sémantiques abstraites (Osgood) :
- Evaluation (négative/positive)
- Pouvoir (potentiel de critique) (faible/fort)
- Activation (passif/actif)
- la théorie cognitive (Sherer): dimensions d'appraisal »: cause/conséquence/nouveauté/valence/etc«.



Annotation des émotions étiquettes verbales

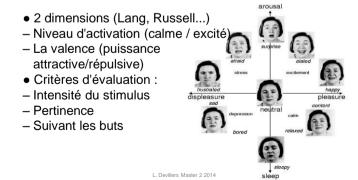
'Word' type intermediate labelling: a menu

You have been reviewing a particular kind of material. How useful would you find each of these words to describe what you saw and/or heard?

	Not at all	Some use	A lot of use	Essential					
admiration	0	- 1	2	3					
affection	0	1	2	3	to				
amusement	0	1	2	3tensio	•	0	1	2	3
annoyance	0	1	2	3 _{trust}		0	1	2	3
anxiety	0	1	2	3 warin	ss	0		2	3
						U	_ '	- 2	3
				wearii	ness	0	1	2	3
				worry		0	1	2	3

L. Devillers Master 2 2014

Modèle multi-dimensions (1)



Choix des étiquettes

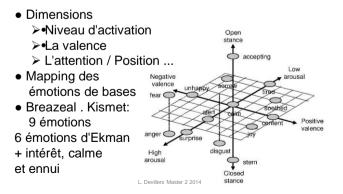
Etude (Cowie): liste des 50 émotions les plus importantes pour un ensemble de 50 personnes testées à partir d'une liste de termes définissant des termes-émotions en anglais.

Etude (Sherer): 1000 personnes interrogées, décrire l'émotion la plus forte ressentie la veille et nommer: ensemble de 40 émotions les plus fréquentes.

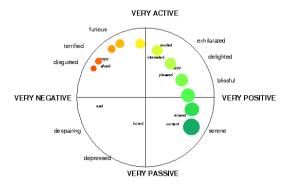
- 2 stratégies devant une nouvelle application:
- étiqueter de façon libre les émotions et annoter le corpus, plusieurs passes nécessaires pour déterminer un sous-ensemble de classes
- sélectionner à partir d'une liste (Cowie ou Sherer par exemple) préétablie des émotions pertinentes pour une tâche donnée, vote majoritaire.

L. Devillers Master 2 2014

Modèle multi-dimensions (2)



Feeltrace (Cowie 2000)



L. Devillers Master 2 2014

Annotation des émotions appraisal-type (2)

'Appraisal' type fine labelling

The broad headings split to give 18 'dimensions'

NO	VELTY	
	Suddenness	High, medium, low
•	Familiarity	High, medium, low
•	Predictability	High, medium, low
	FRINSIC EASANTNESS global	High, medium, low

GOAL SIGNIFICANCE	
Concern Relevance	- person concems (survival, bodily integrity, fulfillment of basic needs, self- esteem) - relationship concerns (establishment, continued existence and intactness of relationships, cohesion of social groups) - social order concerns (sense of orderliness, predictability in a social environment including fairness & appropriateness)
Outcome Probability	High, medium, low
•Relation to expectation	Consonant, dissonant
•Conduciveness to goals	Conducive, obstrucive
•Urgency	High, medium, low

L. Devillers Master 2 2014

Annotation des émotions appraisal-type (Sherer)

'Appraisal' type fine labelling

- Scherer has proposed a scheme based on the idea that emotion is inseparable from the way the organism 'appraises' its situation.
- Some aspects of the theory are debatable, but that does not affect the descriptive power of the scheme.
- · Appraisals are described under 5 broad headings.

NOVELTY
INTRINSIC PLEASANTNESS
GOAL SIGNIFICANCE
COPING POTENTIAL
COMPATIBILITY with STANDARDS

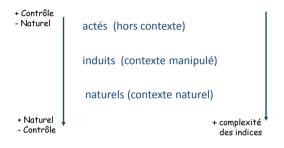
L. Devillers Master 2 2014

Annotation des émotions appraisal-type (3)

'Appraisal' type fine labelling 18 'dimensions'

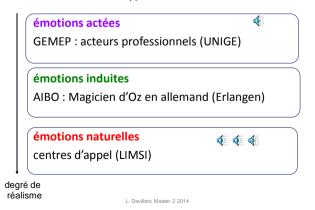
COPING POTENTIAL	
Cause: Agent	Self, other, nature
Cause: Motive	Intent, negligence
Controllability of event	High, medium, low
Power of agent	High, medium, low
Adjustment possible to agent's own goals?	High, medium, low
COMPATIBILITY with STANDARDS	
External (norms or demands of a reference group)	High, medium, low
Internal (self ideal or internalized moral code)	High, medium, low

Typologie des corpus émotionnels



L. Devillers Master 2 2014

Différents types de données



Données actées hors contexte

Colère Peur Tristesse Joie

L. Devillers Master 2 2014

Données naturelles en contexte (ex: centre d'appels)

Avantages

- contenu émotionnel adéquat pour des applications réelles
- enregistrement imperceptible
- émotion exprimée uniquement par la voix (pas d'ambiguités inhérentes à l'expression multimodale des émotions)
- variabilité dans les locuteurs

Inconvénients

- moins bonne qualité des signaux
- contenu émotionnel peu fréquent et non équilibré
- émotions mélangées et peu prononcées
- pré-traitement coûteux
- problèmes éthiques : nécessité d'anonymiser les données, non diffusion

Corpus étudiés - LIMSI-CNRS

Corpus	Audio/ Audio-Visuel	Taille	#locuteurs	Émotions annotées
BOURSE		4.5h	100 appelants/ 4 agents	Colère, Peur, Satisfaction, Excuse
				Colère, Peur, Satisfaction, Excuse
CEMO 🍕				20 classes 7 macro
EmoTaboo	Jeux	7h30	10 locuteurs	> 20 classes
EmoTV	Interviews TV	< 1h	100 locuteurs	14 - 35 classes 7 macro
SAFE	Fiction	7h	400 locuteurs	Peur, autres Neg, Pos

L. Devillers Master 2 2014

Exemple de corpus dans la communauté

Corpus	Content	# All segments	hh:mm	# emotions	# Sub	Rec	AIM
AVIC	English variable		1:47	3 (interest)		spont normal	HH office
DES	Danish fixed		0:28	5 (Ekman)		acted	studio
EMO-DB	German fixed	494	0:22	7 (~Ekman)		acted	studio
eNTERFA CE	NNEnglish fixed	1277		6 (Ekman)	42	acted normal	Students office
SAL	Irish variable			PA		spont normal	WoZ
Smart- Kom	German variable	3823		4 [10]		spont. Noisy	HCI
SUSAS	US fixed			3 [5]		mixed noisy	Noise
VAM	German fixed	946	0:47	PAD	47	spont normal	Broadcasts TV

Corpus naturels étudiés

- 1/ Reading-Leeds database (Greasley 95, Roach 98): interview radio-télévisées expériences émotions fortes, 4.5h, type d'émotions vaste, audio (anglais).
- 2/ Belfast natural database (Douglas-Cowie 00): 298 clips télévisés de 10 à 60 s, discussion, interview scientifiques, 125 locuteurs (31 h, 94 f), type d'émotions vaste, audio-visuel (anglais).
- 3/ France (France 00): session de thérapie et de post-thérapie au téléphone, 115 locuteurs (67 h, 48 f), dépression, état suicidaire et neutre, audio (anglais).
- 4/ Chung (Chung 00): interview télévisées, 6 américains 1 coréen, joie, tristesse, neutre, audio-visuel (coréen, américain)
- 5/ Geneva Airport Lost Luggage (Sherer 97): Video et interview de passagers, 109 locuteurs, colère, bonne humeur, tristesse, indifférence, stress, audio-visuel
- 6/ CREST « parole expressive »: 1000 heures en 5 ans (250 heures, 10% transcrits),
- « Minidisc recorder», qualité > MPEG (Campbell02), pb micro, qualité difficile, conversations + TV, vidéo... (japonais, chinois, anglais)

L. Devillers Master 2 2014

Problématiques

2003 : deux états de l'art sur plus d'une centaine d'études ([Scherer 2003], [Juslin et Laukka 2003]):

"Most of the studies in this area <u>lack theoretical and analytical rigor</u>, and some of the most serious shortcomings are the following: <u>using actor portrayed emotion utterances</u>, as opposed to naturally occurring emotional vocalizations; not systematically controlling important variables such as the number of speakers, the type of emotions studied, the instructions for portrayal, and the verbal material used." [Kappas et al. 1991 p213]

- Etudes s'appuyant sur des données actées souvent par peu d'acteurs
- Peu de classes étudiées
- Pas de validation systématique des étiquettes
- Peu d'indices

Bilan sur les corpus en 2012

Encore très peu de corpus "Real-life" 2012 (Workshop LREC06, 08, 10, 12) - 2014

->résultats obtenus sur des corpus artificiels ne sont pas utilisables pour des corpus spontanés.

Valeur d'un corpus:

- 1/étendu (nb locuteurs, dialectes, genre, types d'émotions,...),
- 2/ authenticité (naturel)
- 3/ contexte (sémantique, structurel-syntaxique, intermodal, temporel)
- 4/ descripteurs (descripteurs micro-prosodiques (ex: ToBi), qualité de voix, indices non verbaux (rires, raclement de gorge)...)
- 5/ accessibilité (format, problème d'éthique, de propriété)

L. Devillers Master 2 2014

Protocole LIMSI

- ✓ Protocole d'annotation itératif:
 - Définition
 - Étiquettes émotion : → Dimensions abstraites : étiquettes dépendant de la tâche
 - dimensions universelles Unités segmentales: toute la séquence, tour de parole, groupe sémantique, mots, etc.
 - - Une étiquette ou une combinaison d'étiquettes
 - Plus des dimensions abstraites par segment
 - Meta-annotation: contexte de la situation, etc
 - - Mesure de fiabilité inter-annotateur
 - Tests perceptifs

Partie 5: Expérience d'annotation de corpus

Ouelle unité annoter? - Comment rendre compte de leur richesse et de la complexité des données? - Contexte ? - Comment former et évaluer des annotateurs experts? - Combien d'annotateurs faut-il? - Comment valider les annotations ?

L. Devillers Master 2 2014

Emotions mélangées dans des contextes réels

- · Perception de mélanges d'émotions (modèle push/pull - Scherer)
 - émotions exprimées et ressenties
 - · Expression mélangée de la colère/désespoir
 - émotions exprimées conflictuelles (indices positifs et négatifs)
 - · Expression d'un soulagement dans les paroles, larmes
 - émotions masquées par les règles sociales, stratégies de communication
 - · Expression positive/ressenti négatif : homme politique
 - Expression négative/ressenti positif : mère de famille gronde son enfant

L. Devillers Master 2 2014 L. Devillers Master 2 2014

Mélange d'émotions positive/négative

• LIMSI – Interviews dans le journal TV

Evènement cause: Mauvais résultats non attendus aux élections Evènement en cours: Interview Rassurer les électeurs

Indices multimodaux: parole Sourire tendu

-Mélange d'émotions:

- Déception masquée par du Courage



Outils d'annotation (vidéo)



L. Devillers Master 2 2014

Outil d'annotation (AUDIO) : transcriber



L. Devillers Master 2 2014

Hiérarchie des étiquettes

Valence	Classe large (7 classes)	Etiquettes fines (20 classes + Neutral)		
	Peur	Peur, Anxiété, Stress, Panique		
Négatif	Colère	Agacement, Impatience, Colère froide Colère chaude Tristesse, Déception, Résignation, Désarroi, Embarras Désespoir		
	Tristesse			
	Douleur	Douleur		
Négatif ou Positive	Surprise	Surprise		
Positive	Positif	Intérêt, Compassion, Amusement, Soulagement		
Neutre	Neutre	Neutre		

Validation des annotations

Pour essayer de pallier à la subjectivité de l'annotation, il est préférable d'utiliser plusieurs annotations en // (5 Batliner, souvent 3 ou 2)

- -> Vote majoritaire (calcul de soft vecteur à partir d'une matrice de confusion entre annotateurs)
- -> Mesure d'intercoder agreement (Kappa) ne veut rien dire ?
- -> Mesure d'entropie entre différentes annotations et l'annotation du système. On peut ainsi juger si le système fait autant d'erreur que les annotateurs!

L. Devillers Master 2 2014

Accord inter-annotateurs : kappa

une composante aléatoire et une composante réelle:

$$\kappa = \frac{P_{obs} - P_{\text{exp}}}{1 - P_{\text{exp}}}$$

 P_{obs} : proportion d'accord observés P_{exp} :proportion d'accord due au hasard

- 1 accord parfait
- 0 pour 2 jugements indépendants
- -1 lorsque les juges sont en total désaccord
- Données boursières : 0.8 (appelant + agent)
- Majeur CEMO : 0.61 (appelant) 0.35 (agent)

L. Devillers Master 2 2014

AUDIO-VISUEL: Annotation du corpus interviews EmoTV



L. Devillers Master 2 2014

Rejouer les annotations manuelles avec un Agent Conversationnel



Intérêt

Partie 6: Indices des émotions dans la Parole

Quels sont les indices les plus pertinents pour discriminer les émotions ?

Est-il possible de tous les obtenir de manière automatique ?

Comment les combiner ?

L. Devillers Master 2 2014

Mélange d'indices



Communication non verbale

Il est estimé que plus de 80% de notre interaction avec les autres se fait par une communication non verbale, ou par le langage du corps.

De plus la majeure partie des indices exprimés sont pilotés par notre subconscient.

Ce sont des indices importants de ce que nous ressentons et pensons

L. Devillers Master 2 2014

Mélange d'indices audio et vidéo



Indices émotionnels

- > Certains indices d'ordre physiologique et expressifs
 - pigmentation de la peau, rythme respiratoire
- Plusieurs indices d'ordre physionomique
 - > traits du visage, regard, posture
- Des indices prosodiques
 - > ton, timbre, rythme de la voix, etc...
- > Tous les indices verbaux possibles
 - indices lexicaux, disfluences (euh, silence, allongement syllabique),
 - Répétitions, mots tronqués, autocorrection
- Tous les indices non verbaux de type :
 - rire, souffle, pleurs, raclement de gorge

L. Devillers Master 2 2014

Les articulateurs

- · le larynx : par ses mouvements souples et verticaux.
- · le voile du palais : en obturant ou pas les cavités nasales supérieures :
- · les mâchoires : plus ou moins ouvertes ;
- · la langue : par son côté polymorphe ;
- · les dents : plus ou moins fermées. plus ou moins obturées par la langue, plus ou moins obturées par les lèvres ;
- · les lèvres : plus ou moins fermées, plus ou moins pincées, plus ou moins étirées.

L. Devillers Master 2 2014

Le modèle source/filtre

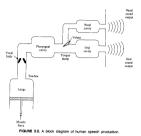
Composants :

- Les poumons (pression d'air)
- 2 éléments qui transforment en un signal source:

- source Les cordes vocales (signaux périodiques)
 - Les constrictions du conduit vocal qui produisent des turbulences (bruit: friction, impulsion)
 - Le conduit vocal



- En partie immobile
- En partie mobile: aussi appelés articulateurs



L. Devillers Master 2 2014

Perceptivement

On peut parler:

plus ou moins fort (Intensité) :

cause : elle est causée par une énergie plus ou moins forte en provenance du diaphragme ;

conséquence : cette énergie provoque une pression plus ou moins forte de l'air sous la glotte ;

résultat sonore : la variation de l'amplitude du son émis et plus ou moins

plus ou moins haut (Hauteur):

cause : périodicité plus ou moins grande du battement des cordes vocales l'une contre l'autre :

conséquence : cette périodicité produit une fréquence de la variation de pression de l'air plus ou moins grande ;

plus ou moins clair (Timbre):

cause : positionnement du voile du palais ;

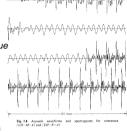
conséquence : le son est plus ou moins riche en harmoniques graves ou aiguës;

Le signal résultat

- Un signal complexe qui change de propriétés constamment:
 - Quelques fois périodiques
 - Quelques fois bruités
 - Approximativement stationnaire sur une fenêtre de ~20 millisecondes
- Et qui contient un grand nombre d'informations
 - Text information linguistique
 - Non verbal information paralinguistique
 - Identité du locuteur
 - Genre
 - Stress, accent
 - Origine sociale
 - Etats mentaux
 - Etat émotionnel

Etc. ...

L. Devillers Master 2 2014



Prosodie

La prosodie est un autre nom pour la partie des informations paralinguistiques, composées de :

- Intonation la voix dont le timbre change au fil du temps
- Intensité changement de l'intensité au fil du temps - Problème : quelques segments sont en soi plus faibles que d'autres
- Rythme débit de parole

La prosodie n'inclut pas la qualité de voix, mais la qualité de voix fait aussi partie des informations paralinguistiques

L. Devillers Master 2 2014

Comment l'information est-elle codée?

Information textuelle-

Principalement dans le filtre

Segment de tour de parole

Information paralinguistique

Principalement dans les paramètres source

- Pression des poumons détermine l'intensité
- Périodicité vocale détermine la fréquence fondamentale
- La configuration de la glotte détermine "la qualité de voix"

L. Devillers Master 2 2014

Extraction de caractéristiques du signal de parole

Etape 1 - Extraction de caractéristiques brutes

Fréquence fondamentale Intensité Qualité de voix Pauses Durée des segments

Etape 2 – Réduction du nombre d'indices

Stylisation

Ex: Modèle sur le pitch

Statistiques sur les signaux -> Calcul de fonctionnelles Pentes

Max, Min, Moyenne, déviation standard Position relative, Range

, ,

Pitch

- Pitch: La fréquence fondamentale
 - Beaucoup de recherches consacrées à la détection du pitch
 - Problème composé de deux sous-problèmes :
 - Pour un signal donné est là la périodicité du tout ?
 - Si c'est ainsi quelle est la fréquence fondamentale ?
 - Facteurs de complexité :
 - · Qualité de voix du locuteur l'enrouement, etc.
 - Bruit, voix superposées, filtres comme dans la téléphonie)
 - Dans le contexte d'émotions :
 - De petites erreurs sont acceptables
 - De grandes erreurs (des sauts d'octave, des fausses détections) sont catastrophiques -> filtres

L. Devillers Master 2 2014

Qualité vocale

Plusieurs mesures sont utilisées pour l'évaluer :

- Irrégularités locales dans le pitch (jitter) et l'intensité (shimmer)
- Proportion entre composantes harmoniques et composantes sonores
- Distribution d'énergie dans le spectre (hautes fréquences)
- Affectée par une multitude de facteurs autres que les émotions
- Mais un facteur important pour détecter les émotions!

L. Devillers Master 2 2014

Intensité

Semble être plus simple que le pitch!

- L'intensité est tout à fait facile de mesurer ...
- Encore plus sous l'influence de facteurs environnants, l'intensité est gravement affectée par :
 - Distance du microphone
 - Bruit de fond
 - Enregistrement d'environnement
- Sans normalisation l'intensité est presque inutile!

L. Devillers Master 2 2014

Segments

- Il y a des façons différentes de définir précisément la notion de segments
- Il y a différentes fenêtres temporelles utiles
- La segmentation automatique est difficile, quoique pas aussi difficile que la reconnaissance de la parole
- Même les frontières de segment peuvent donner des informations importantes, liées au rythme un composant important de la prosodie

Facteurs de complexité

- Bruit de fond
- Superposition de voix
- Variabilité d'élocution des locuteurs
- Variabilité dans la qualité de l'enregistrement

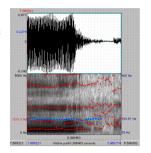
L. Devillers Master 2 2014

Agent : Il doit y avoir un problème Client1 : J'ai peur d'avoir perdu de l'argent, j'ai mal compris sans doute ? Client2 : J'ai j'ai euh peur oh la la d'avoir euh euh perdu, enfin expliquez-moi ? Client3 : oh la la la bon donc euh euh donc oh non, hein dites-moi ? > Dislocation lexicale, syntaxique, sémantique > Dislocation prosodique

L. Devillers Master 2 2014

La prosodie (outil Praat)

- phénomènes d'accentuation et d'intonation avec 4 dimensions
 - la fréquence fondamentale F0 et les formants
 - l'intensité (=énergie)
 - la durée (rythme, silences)
 - la qualité de la voix (tendue, rauque, NAQ...)



L. Devillers Master 2 2014

Catalogue de transformations de voix qualité vocale IRCAM

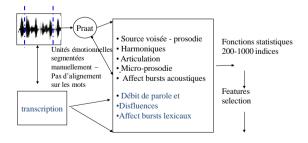


Paramètres paralinguistiques

- Source voisée prosodie (F0, E, rythme)
- Harmoniques (timbre)
- Articulation (formants)
- Micro-prosodie (jitter, shimmer)
- Mfccs

L. Devillers Master 2 2014

Module paralinguistique - détection



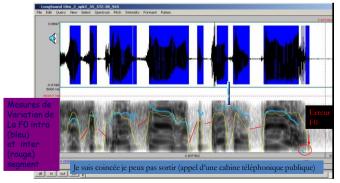
Thèse de L. Vidrascu - Soutenue en Déc. 2007

Marqueurs affectifs non verbaux

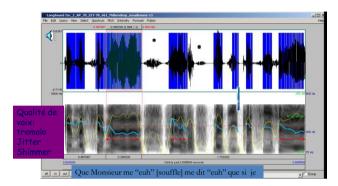
- hésitations (« euh ») : nombre et taille des hésitations
- silences : longueur des silences, nombre de silences
- Toux, rires, pleurs, respiration (peu nombreux)
- Backchannel

L. Devillers Master 2 2014

CEMO: Anxiété/Peur/Stress



CEMO: Désespoir/Désarroi



L. Devillers Master 2 2014

A not-so-old example: (Murray and Arnott, 1993)

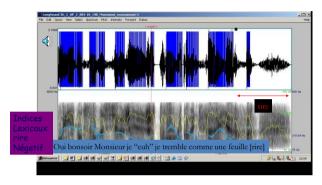
	Anger	Happiness	Sadness	Fear	Disgust
Speech rate	Slightly faster	Faster or slower	Slightly slower	Much faster	Very much slower
Pitch average	Very much higher	Much higher	Slightly slower	Very much higher	Very much lower
Pitch range	Much wider	Much wider	Slightly narrower	Much wider	Slightly wider
Intensity	Higher	Higher	Lower	Normal	Lower
Voice quality	Breathy, chest tone	Breathy, blaring	Resonant	Irregular voicing	Grumbled chest tone
Pitch changes	Abrupt, on stressed syllables	Smooth, upward inflections	Downward inflections	Normal	Wide downward terminal inflections
Articulation	Tense	Normal	Slurring	Precise	Normal

Table 2. Summary of most general correlates to emotions in speech. Table from Murray & Arnott (1993).

- Résultats très qualitatifs
- Présuppose que l'on s'interesse aux émotions primaires

L. Devillers Master 2 2014

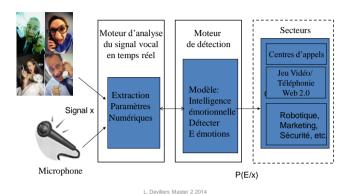
CEMO: Anxiété



L. Devillers Master 2 2014

Partie 7: Classification

Module de détection des émotions LIMSI-CNRS



Peut on reconnaître les émotions?

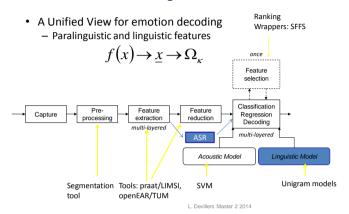
 La machine vise des performances de l'ordre de celles des humains

Les performances humaines (Scherer 1998, 2000, 2003):

- comparaison de 30 études : taux de discrimination plus de 5 fois supérieurs aux taux du hasard.
- 14 émotions jouées par 12 professionnels
 - -> taux de reconnaissance de 48%
- performances humaines autour de 60% pour la reconnaissance de 6 émotions (Colère, Peur, Joie, Tristesse, Dégoût, Etat Neutre) et de 65% en moyenne pour 5 émotions
- Colère: 75%; Peur et Joie: 60%; Dégoût 31%
- étude : auto-annotation vs. jugement perceptif

L. Devillers Master 2 2014

Processing Overview



Modèle de détection des émotions (indices lexicaux)

- Chaque modèle : interpolation entre un modèle unigram d'une émotion et un modèle général calculé sur l'ensemble des émotions
- Détection:

$$\log P(u/E) = \frac{1}{L_u} \sum_{w \in u} tf(w,u) \log \frac{\lambda P(w/E) + (1-\lambda)P(w)}{P(w)}$$

Procédures de prétraitements standards :
 « stemming », « compounding » et « stopping »

Détection de 5 émotions

· Résultats (scores movens)

Total	ANG	FEA	EXC	SAT	NEU
68%	56%	38%	68%	88%	88%

- · Satisfaction marquée lexicalement
 - « D'accord, merci»
- Peur détection au niveau syntaxique
 «ou alors je vends des ou alors je je je vends je ne sais pas encore»
- · Réf: Devillers, Vasilescu, ICME 2003

L. Devillers Master 2 2014

Classificateurs utilisés

- Arbres de décision
 - Arbre simple avec élagage (C4.5)
 - Meta algorithmes : combinaison de modèles :
 - Bagging (ADTree): t modèles avec un ensemble d'apprentissage modifié aléatoirement, vote majoritaire combinant t prédictions.
 - Boosting (Adaboost): poids associé à chaque exemple de l'ensemble d'apprentissage.
- Support Vector machines (SVM)
- Procédures de Jack-Knifing pour l'apprentissage
- Nos résultats expérimentaux sont corrélés avec ceux de la communauté scientifique et montrent peu de différences entre les modèles

L. Devillers Master 2 2014

Items lexicaux pertinents/classe

Colère	Peur	Satisfaction	Excuse
pas_normal	souci	accord	faute
mots d'argot	peur	merci	erreur
énervé	panic	parfait	désolé
gênant	je crains	excellent	excuse

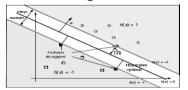
- •Mots pour Peur et Colère : problème, bizarre
- Expressions idiomatiques : Je m'en fous, Je m'assois dessus, je suis marron, je vais perdre ma culotte

L. Devillers Master 2 2014

Les algorithmes d'apprentissage

- Les arbres :
 - chaque nœud correspond à un test et chaque feuille à une classe. La décision de l'attribut à tester se fait par une mesure d'entropie.
- → P
 - un élagage de l'arbre à est nécessaire pour éviter le sur-apprentissage.
- Les méta algorithmes : combiner les sorties de différents modèles pour avoir de meilleurs performances.
 - <u>Bagging</u>: combinaison des prédictions de t modèles par vote majoritaire. Les t modèles sont obtenus en modifiant de manière aléatoire l'ensemble d'apprentissage.
 - ■Boosting : A chaque exemple de l'ensemble d'apprentissage est associé un poids.
 - Au début, tous les poids sont égaux.
 - A chaque itération l'algorithme retourne une hypothèse et les poids des exemples mal classés sont augmentés.
 - L'hypothèse finale est une combinaison des hypothèses de chaque itération, avec un poids plus important pour les hypothèses avec le moins d'erreurs

L es SVM (séparateurs à vaste marge)



Hyperplan optimal de marge 1/||w|| (schéma tiré de l'article de Cornuéjols [SVM])

- Recherche de l'hyperplan optimal séparant les données.
- Dans le cas d'échantillon non linéairement séparables, le coefficient C règle le compromis entre la marge et le nombre d'erreurs admissibles.
- Solution avec des fonctions bilinéaires symétriques positives «fonctions noyaux » choisies par l'utilisateur.

L. Devillers Master 2 2014

La sélection des attributs

- Diminuer le nombre de descripteurs sans nuire à la qualité des résultats
- · Plusieurs méthodes :
 - évaluation d'un attribut / un sous-ensemble d'attributs
 - "Wrappers" (algorithme d'apprentissage/ "filters" (mesures d'entropie)
- Pas de baisse de performance après sélection des attributs
- Performances comparables pour différents algorithmes de selection

Observation

Reconnaissance de la parole

Rec

L. Devillers Master 2 2014

Algorithme et mesures de performances

Score	formule	émotions détéctées
WAR(Recognitio n Rate)	$\frac{\sum m_{ii}}{Total}$	$\begin{bmatrix} m_{11} & m_{12} & \dots & m_{1j} & \dots & m_{1n} \\ \dots & \dots & \dots & \dots \end{bmatrix}$
UAR (Class wise)	$\frac{1}{n}\sum \frac{m_{ii}}{m_i}$	$m_{21} \ m_{22} \ \dots \ m_{2j} \ \dots \ m_{2n}$
FF score	$\frac{RR + CL}{2}$	
Précision (par émotion i)	$\frac{m_{ii}}{m_i^* + m_i}$	$\begin{bmatrix} \dots \\ m_{n1} \ m_{n2} \ \dots \ m_{nj} \ \dots \ m_{nn} \end{bmatrix}$
Recall (par émotion i)	$\frac{m_{ii}}{m_i}$	$m_i = \sum_{i=1}^{n} m_{ij} \qquad m_i^* = \sum_{i=1}^{n} m_{ji}$

L. Devillers Master 2 2014

L. Devillers Master 2 2014

33

Analyse des performances

Influence du rôle : agent/appelant

- ex: Colère/Neutre: colère plus marquée chez les appelants

UAR	Test_Agent	Test_Appelant		
Apprentissage_Agent	82%	77%		
Apprentissage_Appelant	75%	82%		

Influence du genre : homme/femme

- différentes émotions dans des situations similaires [Hess 2006]
- différentes expression d'une « même » émotion
 - pas de différence pour des tâches « simples » (Peur/Neutre ~85%)
 - Tâches complexes: scores meilleurs pour des systèmes entraînés sur des locuteurs du même sexe

L. Devillers Master 2 2014

Détection de 4 émotions (CEMO, appelant)

Modèle lexical

Total	Colère	Peur	Tristesse	Soulagement
78%	59%	90%	34%	86%

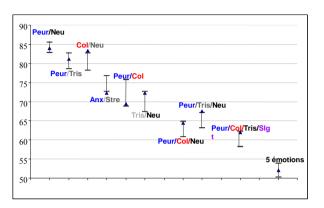
Modèle paralinguistique

Total	Colère	Peur	Tristesse	Soulagement
59.8%	39%	64%	57%	58%

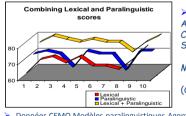
Etat de l'art (Batliner 2005) : < 60% 4 classes indices paralinguistiques

L. Devillers Master 2 2014

Les performances CEMO



Détection 2 classes Négatif/Neutre



> Données BOURSE Appelant + Agent: Combinaison linéaire des Scores (10 ensemble de tests)

Moyenne: 5% gain performances (CHM, Narayanan)

Données CEMO Modèles paralinguistiques Agent: 83% Appelant: 72%

Etat de l'art (Batliner 2005) : < 80% 2 classes indices paralinguistiques

State of the art: Speaker Independent performances

711	Jepei	IUCITI	PCI IV	Ji illuli	CCS		
Corpus	#emotions	Emotion		Activation		Valence	
		UAR	WAR	UAR	WAR	UAR	WAR
AVIC 21 loc.	3 (interest)	56.5	68.6	66.4	76.2	66.4	76.2
DES 4 loc.	5 (Ekman)	59.9	60.1	87.0	87.4	70.6	72.6
EMO-DB 10 loc.	7 (~Ekman)	84.6	85.6	96.8	96.8	87.0	88.1
eNTERFACE 42 loc.	(F)	72.5	72.4	78.1	79-3	78.6	80.2
SAL 4 loc.	Pleasure Activation	29.9	30.6	55.0	55.0	50.0	49.9
Smart-Kom 79 loc.	4 [10]	23.5	39.0	59.1	64.1	53.1	75.6
SUSAS 7 loc.	3 [5]	61.4	56.5	63.7	77-3	67.7	68.3
VAM 47 loc.	PAD	37.6	65.0	72.4	72.4	48.1	85.4
Моу.	-	53-5	59-9	71.1	75-4	64.5	68.3
L. Devillers Master 2 2014							



Questions?

devil@limsi.fr