

---

# Expression faciale d'émotions par un narrateur virtuel pour des enfants

Nesrine Fourati<sup>1</sup>, Adeline Richard<sup>2</sup>, Nicolas Sabouret<sup>1</sup>,  
Jean-Claude Martin<sup>1</sup>, Emilie Chanoni<sup>2</sup>, Céline Clavel<sup>1</sup>

1. LIMSI, CNRS, Univ. Paris-Sud, Université Paris-Saclay,  
Bat 508, rue von Neuman, 91405 Orsay Cedex, France  
{fourati, nicolas.sabouret, jean-claude.martin, celine.clavel}@limsi.fr

2. PSY-NCA, Université de Rouen, 76000 Rouen, France  
{adeline.richard, emilie.chanoni}@etu.univ-rouen.fr

---

**RÉSUMÉ.** Cet article explore comment la dimension affective d'une histoire peut être prise en compte dans le cadre d'une narration réalisée par un agent virtuel et quel est son apport pour la compréhension. Nous proposons un modèle computationnel qui s'appuie sur le modèle d'évaluation cognitive des émotions (appraisal) proposé par Scherer (2001) pour doter un agent virtuel de la capacité d'exprimer des émotions en fonction des différentes situations narratives. Nous avons implémenté ce modèle dans un agent narrateur qui a raconté une histoire à vingt-trois enfants âgés de six à dix ans. Les résultats de cette étude montrent que la compréhension générale de l'histoire n'est pas affectée par les capacités expressives de l'agent. Néanmoins l'expressivité émotionnelle de l'agent semble orienter le point de vue adopté par les enfants lors de leur compréhension des événements émotionnels.

**ABSTRACT.** This paper focuses on the development of an expressive virtual storyteller for children and raises the challenges encountered to endow a virtual agent with storytelling capabilities. We focus on facial expressions of emotions and we present our computational model based on an appraisal theory of emotion. We also discuss different viewpoints a storyteller can adopt to evaluate the expressive content of a story event. We describe a preliminary study that we conducted with 23 children. We discuss the impact of facial expressions of appraisals on children's memorization of story events, their perception of characters' appraisals, their subjective perception of the virtual storyteller and more generally how emotion combines with joint attention.

**MOTS-CLÉS :** narrateur virtuel, expression faciale, émotion.

**KEYWORDS:** expressive virtual storyteller, facial expression, appraisal.

---

DOI:10.3166/RIA.31.537-556 © 2017 Lavoisier

## 1. Introduction

### 1.1. *Le récit narratif dans le développement de l'enfant*

Les « histoires », c'est-à-dire les récits narratifs, jouent un rôle important dans le développement des enfants. Selon Bruner (1996, 2000), les récits peuvent être considérés comme un média favorisant l'élaboration de connaissances sur le monde permettant ainsi d'insérer l'enfant, et plus généralement l'individu, dans la culture. Les récits véhiculent les normes et les valeurs de celle-ci. Ils racontent des situations variées, des réactions diverses. Ainsi, ils informent l'enfant des moyens dont il dispose pour réagir, le confrontent à d'autres réalités possibles que la sienne, l'initient à ce qui est défini comme « normal » ou pas, et aiguïssent de cette façon sa sensibilité à la norme. Les récits permettent une rencontre virtuelle à toute sorte de situations qu'il faudrait des années, voire plusieurs vies, pour expérimenter. Grâce au récit, l'individu n'a pas besoin de vivre toutes ces situations pour en comprendre les ressorts. Le récit propose un schéma de la situation facilitant son appropriation. De cette manière, l'individu peut se confronter à une foule de personnes, toutes différentes, toutes vivant des situations distinctes qu'il n'a pas nécessairement lui-même vécues et tout de même comprendre et partager leurs enjeux.

Le récit permet également à l'enfant de développer sa propre réflexion par rapport aux événements. Il favorise leur mise à distance. En les verbalisant ou en se les représentant, l'enfant et plus généralement l'individu, adopte une position plus distanciée par rapport à la réalité. Cela peut faciliter la mise en perspective de la situation et la compréhension des réactions propres et de celles d'autrui. Le récit contribue de cette manière à développer le modèle de soi, aidant l'enfant à mieux se connaître (Bruner, 1990).

D'une certaine manière, le récit concourt à un plus grand éclairage sur les émotions humaines. Selon Oatley (2004), il permet d'éprouver des émotions variées, de développer des connaissances à leur sujet. Il propose des cadres de référence qui informent sur les différentes dimensions des émotions, sur ce qui est susceptible de les occasionner, sur leurs manifestations... Ces cadres de référence permettent également à l'enfant d'intégrer les normes culturelles d'expressions des émotions. Enfin, le récit permet également à l'enfant de comprendre les émotions selon différentes perspectives. Il favoriserait le développement de théories fonctionnelles sur les émotions, et serait une source d'enrichissement de connaissances à leur sujet.

Il est important de noter que les connaissances sur les émotions s'élaborent à partir de l'expérience de l'individu. Plus la personne expérimente des émotions, plus celle-ci développe des scripts émotionnels c'est-à-dire des stocks de connaissances répertoriant les différentes caractéristiques des émotions. Selon Rimé (1997), chaque émotion éprouvée aura par la suite un impact cognitif du fait qu'elle occasionne la mise en mémoire de tous les éléments présents alimentant ainsi une « base de données » sur les situations émotionnelles, sur leurs déclencheurs et sur les réponses qui s'y produisent. Un savoir émotionnel se constitue de la sorte. D'autres sources vont l'alimenter comme l'interaction sociale. Cette « base de données » s'enrichit tout au long de la vie. À mesure qu'elle s'étoffe, son rôle médiateur dans le

processus émotionnel se développe également. Ce savoir émotionnel est intégré à l'évaluation conduisant à l'émergence ou non d'une réponse émotionnelle et participe à la forme particulière que prendra la réponse émotionnelle éventuelle. Frijda (1993) qualifie de scripts émotionnels les structures qui incarnent des types d'événements et leurs affects associés, les causes possibles et les conséquences des émotions associées et leur signification sociale et personnelle. Ainsi les connaissances sur les émotions ont un rôle primordial puisqu'elles déterminent en partie l'expérience émotionnelle du sujet. Raconter des histoires aux enfants permet donc de les solliciter sur le traitement d'informations émotionnelles et d'enrichir ainsi leurs scripts. Les récits semblent être des outils très utiles pour aider l'enfant à développer ses connaissances sur le monde et plus particulièrement sur les émotions.

### **1.2. Les outils informatiques pour le récit narratif**

Les médias interactifs soutenant le développement de la littératie chez l'enfant sont en plein essor (Noorhidawati *et al.*, 2015). Ces dispositifs sont considérés comme notamment favorisant l'engagement de l'utilisateur (Szafir et Mutlu, 2012), rendant l'activité plus attrayante (Bickmore, *et al.*, 2016) et agréable pour les jeunes enfants (Oberlander et Robertson, 2002). Il semble également que ces nouveaux supports favorisent l'engagement dans la lecture (de Jong et Bus, 2004 ; Moody, *et al.*, 2010), la participation à la communication (Moody *et al.*, 2010), ne perturbent pas la mémorisation des histoires (Ricci et Beal, 2002), voire améliorent la compréhension (Doty *et al.*, 2001).

Néanmoins la question de l'incarnation du narrateur se pose. Ainsi, on observe depuis les années 1990, un envol des recherches sur le développement de robots narrateurs (Mutlu *et al.*, 2006 ; Chella *et al.*, 2008 ; Kory et Breazeal, 2014). Concernant le développement des agents virtuels, un vif intérêt a été porté sur les agents pédagogiques en soulignant leur apport pour les apprentissages en renforçant notamment l'engagement des élèves (Lester *et al.*, 1997, Atkinson, 2002). Les agents virtuels sont également considérés comme des médias intéressants pour introduire la dimension affective de l'interaction. Guo *et al.*, (2015) montrent notamment que des agents conversationnels dotés d'expressions affectives favorisent la motivation des élèves ainsi que leur plaisir d'apprendre comparé à des situations d'apprentissage où l'agent présent n'exprime pas d'affect, voire est tout simplement absent. Un certain nombre d'agents narrateurs ont également été développés. Par exemple Silva *et al.*, (2001) proposent l'agent Papous dont l'objectif est de raconter des histoires de façon expressive et crédible à la manière d'un humain. Papous dispose de deux modules d'expression affective : l'un pour le discours, l'autre pour le corps, et son comportement s'adapte aux choix de l'utilisateur.

Néanmoins, les évaluations concernant l'apport d'un tel dispositif concernent l'ergonomie (Silva *et al.*, 2003) ou interrogent la crédibilité des modalités d'expression en comparant les performances d'un narrateur humain à celles d'un narrateur virtuel (Raimundo *et al.*, 2007). De même, Bleackley *et al.*, (2010) proposent un agent narrateur qui exprime des émotions à travers ses expressions

faciales et corporelles au cours de l'interaction afin d'induire des émotions chez l'utilisateur. L'évaluation du dispositif est réalisée auprès d'adultes et porte sur le changement d'état émotionnel des utilisateurs au cours de l'interaction en comparant la présence ou non de l'agent expressif. Peu d'études questionnent le réel impact des agents virtuels affectifs narrateurs sur la compréhension que les enfants ont de l'histoire et la dynamique émotionnelle véhiculée par l'agent.

### **1.3. Problématique**

Dans cet article, nous discutons des principaux défis que nous rencontrons pour développer un conteur virtuel expressif pour enfants. Nous présentons quelques solutions et nous étudions l'impact de l'expression d'émotions sur la compréhension, par des enfants, des dimensions narratives et affectives du récit.

Nous nous concentrons sur l'expression faciale d'émotions au cours de l'interaction narrative. La plupart des recherches sur l'expressivité émotionnelle des agents narrateurs s'appuient sur les approches catégorielles proposées par Ekman (Bleackley *et al.*, 2010 ; Ochs *et al.*, 2010) et rendent compte des émotions qui sont contenues dans le texte (Bleackley *et al.*, 2010). À la différence de ces travaux, notre proposition repose sur une approche cognitive de l'émotion défendue par Scherer (2001). Ainsi, les expressions émotionnelles de l'agent résultent de l'évaluation des événements de l'histoire et de l'interaction avec l'enfant.

Nous présentons également le cadre dans lequel nous inscrivons notre récit interactif qui repose sur une approche d'inspiration socio-constructiviste notamment défendue par Vygotski (1967) ou encore par Bruner (1983). Nous nous intéressons en particulier à l'initiation de l'*attention conjointe* à partir des comportements verbaux et non-verbaux du conteur virtuel expressif.

### **1.4. Attention conjointe et expression d'émotions**

L'attention conjointe opère quand deux individus ont le même objet d'attention et construisent un sens commun (Crahay, 1999). L'attention conjointe peut être initiée par un des partenaires à partir de ses expressions verbales (par exemple, « regarder ce sujet ») ou de ses comportements non verbaux (par exemple, direction du regard, gestes de pointage).

Plusieurs études en informatique affective mettent en évidence l'importance des cadres de récits interactifs favorisant la participation des enfants à l'histoire racontée par un robot (Fridin, 2014) et ou encore encourageant leur créativité pendant la narration (Theune *et al.*, 2013). Il s'agit à partir de cette approche de rendre actif l'apprenant au cours de ses apprentissages en augmentant notamment le plaisir d'écouter des histoires (Fridin, 2014).

Ainsi, alors que l'expression de l'émotion et l'initiation de l'attention conjointe semblent être deux exigences majeures pour la conception d'un conteur virtuel, seuls quelques systèmes ont été conçus pour approcher conjointement expressivité

émotionnelle et attention conjointe (Adam et Cavedon, 2013 ; Bosse *et al.*, 2007). C'est pourquoi notre travail s'inscrit dans cette perspective et vise à étudier et concevoir un narrateur virtuel expressif qui sollicite l'enfant au cours de la narration notamment sur les séquences émotionnelles. Toutefois, à la différence des travaux précédents, notre modèle informatique ne s'intéresse pas à la modification de la narration mais à l'expression non-verbale d'émotions.

### **1.5. Modèles informatiques de narrateur**

Les premiers modèles de narration faisant intervenir des agents virtuels peuvent être catégorisés en deux approches. D'un part, les modèles de narration interactive (Cavazza, 2002) dans lesquels l'agent est un acteur de la narration, à la manière d'un jeu vidéo interactif. D'autre part, les modèles de narrateur qui cherchent à imiter un humain qui « raconte une histoire » (Silva *et al.*, 2003). C'est dans cette deuxième catégorie que se placent nos travaux.

Dans ce domaine, il existe plusieurs problématiques : de nombreux travaux en Intelligence Artificielle se sont intéressés à la production automatique de narration. (Callaway et Lester proposent une revue de littérature intéressante dans (Callaway, 2002)). On peut distinguer des approches à base de planification (Riedl, 2004) ou à base de simulation (Swartjes, 2008). En s'appuyant sur ces méthodes d'IA, quelques travaux en *affective computing* se sont intéressés à la modulation d'une histoire de manière à susciter des émotions en simulant l'évaluation cognitive des événements (Silva *et al.*, 2003) ou en intégrant des connaissances (Miletitch *et al.*, 2012).

Notre objectif est complémentaire à ces approches puisque nous ne cherchons pas à synthétiser de la narration. Nous partons au contraire d'une histoire déjà écrite que notre agent virtuel raconte à l'enfant. Ce qui nous intéresse, c'est d'étudier comment l'expression d'émotion par l'agent virtuel impacte la perception de la narration par l'enfant.

Le premier problème posé est alors la définition d'un mécanisme de contrôle de l'agent virtuel pour l'expression d'émotions et la gestion de l'attention conjointe. L'originalité du modèle que nous décrivons ci-après est de se baser sur une approche cognitive des émotions défendue par Scherer (2001) mais aussi de prendre en compte le point de vue du narrateur, qui change l'interprétation des événements. L'apport de la théorie de K. Scherer est de considérer l'expression d'émotion comme un élément du processus d'évaluation cognitive et non comme une fonction des émotions (celles-ci n'étant que la désignation d'une résultante physiologique du mécanisme d'évaluation). Notre agent exprime donc directement une évaluation cognitive de la situation, soit par rapport à l'histoire, soit par rapport à la situation d'interaction. Notre objectif est alors de comprendre comment l'observateur (dans notre cas, l'enfant) relie cette expression d'émotion à l'histoire qu'il est en train d'écouter, située dans le contexte d'interaction avec l'agent. Notre propos n'est donc pas simplement de déterminer si le fait d'exprimer des émotions modifie l'attention, mais bien d'étudier comment l'expression du processus d'évaluation cognitive modifie la compréhension de la situation chez l'enfant.

### 1.6. Plan de l'article

Dans la section suivante, nous présentons notre méthodologie pour construire un conteur artificiel capable d'exprimer des émotions et d'initier une attention conjointe avec l'enfant. Il est basé sur l'annotation du scénario d'une histoire pour la génération conjointe de narration expressive et comportements d'initiation à l'attention conjointe. Nous expliquons notamment comment le comportement est contrôlé. La section 3 présente une première étude exploratoire visant à étudier l'impact de l'association d'expressions faciales émotionnelles exprimées par l'agent narrateur durant les phases d'attention conjointe sur la compréhension de l'histoire d'une part, et la perception générale de l'agent d'autre part, dans le cadre d'une interaction narrative réalisée auprès d'enfants âgés de 6 à 8 ans.

Enfin, la dernière section conclut et discute des enjeux et des défis qui restent à mener pour développer un agent narratif autonome et de l'intérêt pour l'interaction narrative de doter l'agent d'une expressivité émotionnelle.

## 2. Méthodologie

Notre narrateur expressif est implémenté dans la plateforme d'agent virtuel MARC (*Multimodal Affective and Reactive Characters*) (Courgeon, *et al.*, 2008). Nous utilisons l'agent féminin de la plateforme, appelé Mary, qui raconte l'histoire en s'appuyant sur les illustrations affichées à côté, comme nous le voyons sur la figure 1. Cet agent a été évaluée de manière informelle par l'équipe de psychologues associée au projet comme étant bien adapté pour des interactions avec des enfants car moins impressionnant que sa version masculine. Comme nous le verrons plus loin, nous avons renforcé son aspect naturel avec un sourire en expression neutre.



Figure 1. Notre narrateur expressif Mary combine l'affichage d'expressions faciales associées à une évaluation cognitive et des gestes déictiques

L'histoire que nous avons choisie est tirée du livre « le ballon perché » de Catherine de Lasa (Auteur) et Robin Gindre (Illustrateur). Dans cette histoire, trois enfants jouent au ballon avant le début des cours. Le ballon se retrouve coincé sur le toit de l'école et les enfants vont essayer de le faire tomber en lançant leurs affaires sur le toit. Mais la situation ne fait que s'empirer puisque toutes leurs affaires se retrouvent à leur tour, coincées sur le toit. Les enfants doivent alors retourner en classe et trouver des astuces pour que le maître ne remarque pas qu'il leur manque un cartable, une trousse, une chaussure... À la fin, une tempête fait retomber les affaires des enfants dans la cour avec le ballon, à la grande surprise de tous les protagonistes.

Cette histoire présente plusieurs intérêts du point de vue de la narration affective : elle se déroule dans un cadre familier et réaliste pour les enfants, mais elle fait apparaître plusieurs situations inattendues pour de jeunes enfants.

Notre méthodologie fonctionne en deux étapes. La première consiste à annoter manuellement le contenu de l'histoire avec des marqueurs qui permettront de contrôler le comportement non-verbal de l'agent virtuel. La deuxième étape est l'animation en temps réel du narrateur virtuel en fonction du contenu de l'histoire et des marqueurs de comportement non-verbal. Notre objectif étant de munir le narrateur virtuel d'une capacité à mener une narration expressive interactive, le contenu initial de l'histoire a été revu par des chercheurs en psychologie pour y ajouter des dialogues interactifs. Cela contribue à établir une activité d'interaction narrative, à mettre en évidence le contenu émotionnel dans les événements de l'histoire et à maintenir l'attention de l'enfant tout au long de la narration. L'histoire modifiée comprend ainsi 160 phrases, réparties en deux catégories : 56,25 % de « phrases narratives », correspondant à l'histoire originale, et 43,75 % de « phrases interactives » permettant de contrôler et de maintenir l'engagement de l'enfant dans la tâche de narration. Nous avons identifié plusieurs types de phrases interactives :

- les phrases interrogatives, dont l'objectif est directement de maintenir l'engagement de l'enfant en l'interpellant sur les événements de l'histoire. Par exemple : « Est-ce que tu crois qu'ils vont réussir à faire tomber le ballon ? » ;

- les phrases affirmatives qui sont utilisées dans l'introduction et dans la conclusion (par exemple : « Bonjour, je m'appelle Mary ») ou pour gérer la transition entre la réponse de l'enfant et une phrase narrative (par exemple : « Je suis d'accord avec toi ; regardons ce qu'il va se passer ») ;

- les phrases d'attention conjointe (ou JA pour joint attention) pour guider l'attention de l'enfant vers des événements ou des objets importants de l'histoire. Par exemple : « Regarde le maître : il a l'air en colère ! »).

Les phrases d'attention conjointe (JA) sont systématiquement associées à des comportements non-verbaux. Tout d'abord, la tête de l'agent et le regard sont orientés vers l'image. Ensuite, un mouvement du torse ou un geste de pointage sont utilisés pour attirer l'attention de l'enfant sur le personnage, l'objet ou l'événement mentionné. Le mouvement du torse est associé aux événements, les gestes de pointage pour les personnages ou les objets. La position de pointage est définie manuellement pour chaque phrase. La synchronisation entre les différentes

modalités est la suivante : le mouvement du regard commence avec le début de la phrase. Il est immédiatement suivi du mouvement du torse ou du geste de pointage. Le mouvement du torse se termine en revenant à une position de repos alors que le geste de pointage est maintenu jusqu'à la phrase suivante, comme illustré sur la figure 1. Après chaque phrase, le narrateur maintient une pause de 400ms, sauf pour les phrases interrogatives où le système attend la fin du tour de parole<sup>1</sup>.

Nous avons aussi produit 160 fichiers audio pour les 160 phrases. Nos expérimentations préliminaires ont montré qu'il n'était pas possible d'utiliser des voix synthétiques avec des enfants : ils étaient gênés par l'aspect non-naturel de la voix (Serban *et al.*, 2013). Nous avons donc demandé à une étudiante en psychologie d'enregistrer la partie verbale de chaque phrase de manière spontanée et expressive. Nous avons synchronisé le mouvement des lèvres avec cette narration en utilisant l'outil de synthèse vocale JSAPI<sup>2</sup> : le contenu des phrases est utilisé pour contrôler les mouvements des lèvres.

Comme nous le verrons dans la section suivante, les différentes phrases peuvent être associées à des expressions faciales d'émotions à partir de l'évaluation cognitive de la situation. Nous ne nous sommes pas intéressés à l'expression d'émotion à l'aide de mouvements du corps : l'agent exprime ses émotions uniquement à travers des expressions faciales qui sont construites de manière cohérente avec le contenu de l'histoire et la prosodie (obtenue de manière spontanée).

### 2.1. *Narration expressive*

Plusieurs théories ont été proposées pour modéliser les caractéristiques morphologiques et dynamiques des expressions faciales d'émotions. Elles sont généralement classées en 3 groupes : les approches catégorielles, les approches dimensionnelles et les approches basées sur l'évaluation cognitive (ou appraisal). Les approches basées sur l'évaluation cognitive ont fortement influencé les modèles conçus en informatique affective (Marsella *et al.*, 2010) et on les retrouve dans de nombreux modèles informatiques d'émotions (une liste de ces modèles est donnée dans (Marsalla et Gratch, 2009)). Dans nos travaux, nous utilisons le modèle CPM (Component Process Model), de Klaus Scherer (Scherer, 2001). Cette théorie de l'évaluation cognitive repose sur l'évaluation en séquence de plusieurs variables qui permet d'évaluer et de décrire une réaction d'adaptation à un événement de l'histoire. De plus, une relation entre les variables d'appraisal et les AU du FACS (Facial Action Coding System) a été proposée par Scherer (Scherer, 2001) et implémentée dans la plate-forme MARC (Courgeon *et al.*, 2014). Par exemple, un événement surprenant implique les AU suivantes : AU1 (lever l'intérieur des

---

1. Dans la version actuelle, la fin du tour de parole est gérée par l'expérimentateur suivant une approche de type « Magicien d'Oz ». À terme, nous voudrions le remplacer par un mécanisme automatique.

2 <http://jsapi.sourceforge.net/>



sourcils), AU2 (lever l'extérieur des sourcils), AU5 (lever la paupière supérieure) et AU26 (baisser la mâchoire).

Nous avons extrait cinq variables d'appraisal (ou *Stimulus Evaluation Checks*) du modèle initial de CPM pour notre étude : la nouveauté (*novelty*), le plaisir intrinsèque (*intrinsic pleasantness*), la contribution aux buts (*goal/need conduciveness*), le contrôle (*coping control*) et le pouvoir d'adaptation (*coping power*). La variable de compatibilité aux normes (*norm/self compatibility*) n'a pas été retenue en raison du peu d'information sur l'impact de cette variable sur les Action Units (AU) d'expression faciale.

Dans notre modèle, chaque variable peut prendre une valeur dans l'ensemble  $\{-1, -0,5, 0, 0,5, 1\}$ . Nous avons annoté manuellement chacune des 160 phrases pour leur attribuer une valeur dans l'espace de variables de CPM : cela définit, pour chaque phrase, l'expression faciale de l'agent narrateur. Cette annotation a été faite par cinq des auteurs de l'article (chercheurs en psychologie et en informatique affective) pour l'ensemble des phrases en tenant compte non seulement du contenu sémantique mais aussi de l'expression d'émotions dans la voix enregistrée, de manière à garantir une cohérence entre les expressions faciales et le discours de l'agent. Nous avons ensuite confronté nos évaluations jusqu'à atteindre un accord pour chaque phrase. La principale difficulté est que l'évaluation de chaque variable nécessite d'avoir une définition claire du point de vue du narrateur. Le même événement peut en effet être perçu comme plaisant ou non selon le point de vue adopté. Par exemple, considérons le moment où le ballon se retrouve coincé sur le toit de l'école. Cet événement est surprenant et décevant du point de vue des personnages de l'histoire mais il est plutôt attendu (d'après le titre de l'histoire) et amusant du point de vue d'un narrateur externe. Il faut donc, pour chaque phrase, se mettre d'accord sur le point de vue adopté par le narrateur. Nous avons ainsi identifié trois points de vue possibles :

- le point de vue « narrateur », qui reste bienveillant mais évalue les événements par rapport à sa propre représentation du monde. Par exemple, lorsqu'à la fin de l'histoire, un enfant embrasse sa botte qu'il est heureux de retrouver, le narrateur est surpris et dégoûté (ce qui n'est visiblement pas le cas du personnage) ;

- le point de vue « narrateur empathique », qui anticipe les émotions des personnages en fonction de leurs attitudes et des événements futurs. Par exemple, lorsque le narrateur se demande comment les enfants vont faire pour gérer la disparition de leurs affaires (phrase de type affirmatif : « je me demande comment ils vont bien pouvoir faire »), il exprime du déplaisir et un empêchement des buts car il sait que la classe va commencer ;

- le point de vue « personnage », qui est utilisé sur des phrases narratives pour traduire les émotions des personnages. Par exemple, lorsque les enfants décident de jouer au ballon (« Qui veut jouer au ballon avec moi ? Moi ! Moi ! »).

Nous avons attribué un point de vue à chaque phrase en fonction de son contenu et nous avons tenu compte de ce point de vue dans l'évaluation cognitive de la phrase. Bien sûr, toutes les phrases ne sont pas porteuses d'émotions : à la fin, 57,50 % des phrases ont été associées à au moins une variable d'appraisal non nulle

(73,91 % de phrases narratives, 6,52 % de phrases interrogatives ou affirmatives et 19,57 % de phrases d'attention conjointe). Sur ces phrases « expressives », 32,61 % adoptent le point de vue « narrateur », 45,65 % adoptent le point de vue « personnage » et 21,74 % le point de vue « narrateur empathique ». En effet, les émotions sont principalement évaluées en fonction du contenu de l'histoire (phrases narratives) et du point de vue des personnages.

Il est important de noter que nous n'avons pas annoté des émotions mais bien des variables d'évaluation cognitives (qui, à terme, pourraient être annotées automatiquement). Dans la section 3, nous verrons comment les expressions faciales résultant de ces évaluations, c'est-à-dire les émotions jouées par l'agent, agissent sur la compréhension de l'histoire par l'enfant.

## 2.2. *Coordination verbale et non-verbale*

La coordination entre les comportements verbaux et non-verbaux peut avoir un impact important sur la crédibilité d'un agent virtuel (Martin *et al.*, 2011). Nous avons travaillé deux aspects de cette question pour notre narrateur affectif : quand exprimer les émotions et comment gérer les expressions neutres.

Des travaux précédents ont montré que la relation temporelle entre l'expression faciale et la parole modifie la reconnaissance, l'intensité perçue et le réalisme des émotions d'un agent virtuel (Buisine *et al.*, 2009). En particulier, les auteurs ont montré que l'expression faciale est perçue comme plus intense et plus réaliste lorsqu'elle couvre tout l'intervalle temporel de parole. C'est pourquoi nous avons adopté cette temporalité dans notre agent virtuel.

Pour les phrases non expressives, il a été montré dans (Carrera-Levillain et Fernandez-Dols, 1994) que les expressions neutres peuvent être perçues comme des états affectifs négatifs. Or nous considérons qu'il est important, dans une interaction avec des enfants, de créer une atmosphère chaleureuse. De plus, notre agent doit apparaître comme bienveillant et amical. C'est pourquoi nous avons configuré Mary de manière à ce qu'elle affiche un « sourire amusé » dans l'état neutre en activant les AU6 (lever les pommettes) et AU12 (tirer le coin des lèvres), comme proposé par (Ochs *et al.*, 2010). En effet, dans cette étude, le sourire amusé est celui qui est considéré comme le plus approprié pour un agent bienveillant et amical. Nous avons aussi ajouté l'AU1 (lever l'intérieur des sourcils) car la configuration neutre des sourcils de notre agent ne paraissait pas naturelle lorsqu'elle est couplée avec un sourire. Enfin, nous avons utilisé AU25 (ouvrir les lèvres) pour alterner entre deux types de sourire au cours de la narration. Ces expressions sont jouées pendant les phrases neutres (c'est-à-dire pour lesquelles toutes les variables CPM sont évaluées à 0). Pendant les pauses de 400ms entre les phrases (sauf pour les phrases interrogatives qui attendent une réponse de l'enfant), l'expression précédente est conservée (l'agent ne revient pas à un état neutre).

Dans la prochaine section, nous présentons notre étude réalisée auprès d'enfants de primaire pour valider l'impact du comportement affectif de l'agent sur la perception de la narration.

### 3. Étude exploratoire

Nous avons conduit une première étude exploratoire afin d'étudier d'une part l'apport de l'expressivité faciale émotionnelle de l'agent à la compréhension de l'histoire et la perception générale de l'agent et interroger d'autre part l'apport de l'expression des différentes variables d'*appraisal* pour la compréhension de la dimension affective de différents événements se déroulant au cours de l'histoire.

#### 3.1. Méthode

##### 3.1.1. Participants

Vingt-trois enfants (dont dix filles) contactés via un centre de loisirs ont participé à notre expérience pendant les vacances scolaires de printemps. Les enfants avaient entre 6 et 10 ans (moyenne = 8 ans ; std = 0,94). Les parents de chaque enfant ont donné leur accord de consentement éclairé et ont convenu que les réponses de leurs enfants à l'étude ainsi que leurs vidéos pourraient être utilisées et publiées à des fins de recherche.

##### 3.1.2. Matériel et procédures

Après une brève présentation collective de la recherche dans le réfectoire du centre aéré, les enfants volontaires pour participer à l'étude ont été invités à se manifester. Individuellement, ils ont ensuite été conduits dans une salle au calme et installés assis devant un ordinateur à une distance d'environ 0,6 mètre d'un moniteur. L'expérimentateur rappelle brièvement le déroulement de la session, lance l'application et se met en retrait dans la salle pour laisser l'enfant seul interagir avec le dispositif.

L'agent narrateur initie l'interaction en posant quelques questions personnelles à l'enfant (son nom et son âge, qui ne seront jamais utilisés par l'agent par la suite) puis lui propose de lui raconter l'histoire du ballon perché. Le récit de l'histoire dure environ 10 minutes. Pendant toute la séance d'enregistrement, l'enfant était enregistré via une caméra et tous les événements de l'histoire ainsi que tous les comportements de l'agent étaient enregistrés.

Douze enfants ont été invités à écouter l'histoire racontée par notre agent virtuel dans l'état expressif, tandis que onze enfants ont été invités à écouter l'histoire dans l'état non expressif. Dans l'état expressif, notre agent virtuel a affiché des expressions faciales correspondant aux différentes variables d'*appraisal* pour chaque énoncé expressif basé sur l'annotation décrite dans la section 3.1. Dans l'état non expressif, notre agent virtuel a montré un sourire amusé tout au long de la tâche de narration (voir section 3.1). L'enfant était libre de faire des commentaires oraux pendant toute la narration mais ils ne l'ont jamais fait hors des sollicitations de l'agent. À la fin de l'histoire, l'enfant a été invité à remplir, avec l'aide d'un adulte, un questionnaire à questions ouvertes portant sur sa compréhension générale de l'histoire et sur sa compréhension plus spécifique de différents événements affectifs se déroulant au cours de l'histoire. Enfin la dernière partie du questionnaire portait

sur son évaluation de l'agent en reprenant certaines sections du questionnaire du Godspeed (Bartneck *et al.*, 2009) notamment sur l'appréciation de l'agent, l'anthropomorphisme, l'intelligence perçue et sur son expressivité.

Le questionnaire de compréhension générale portait sur 16 questions dont les éléments de réponse étaient soit présents durant la narration (10 questions) soit à inférer (6 questions). Les questions portaient soit sur des éléments factuels (8 questions de « mémorisation ») soit sur les états mentaux des personnages (8 questions de « compréhension ») et enfin certaines questions portaient sur des éléments où l'attention conjointe avait été initiée par la narrateur (8 questions) alors que d'autres non (8 questions). Les différentes combinaisons de ces 3 axes (élément présenté à ou inférer, mémorisation ou compréhension et attention conjointe initiée par le narrateur ou non) étaient toutes représentées dans le questionnaire.

Le questionnaire de compréhension des événements émotionnels de l'histoire est une adaptation du questionnaire genevois d'*appraisal* (QGA) développé par Scherer (2001) et interroge spécifiquement certaines variables d'*appraisal*, c'est-à-dire la nouveauté, le plaisir intrinsèque, la contribution aux buts, le contrôle et les normes. Le vocabulaire utilisé a été adapté à un public « jeunesse ». Pour chaque événement, une illustration de l'histoire était présentée à l'enfant et interrogeait le point de vue d'un des protagonistes de l'histoire (pour 4 événements) et le point de vue de l'agent (pour 4 événements). A chaque fois, les 5 dimensions d'évaluation étaient étudiées<sup>3</sup> à partir de propositions pour lesquelles l'enfant devait donner son niveau d'accord via une échelle de Likert en 5 points (par exemple, « *D'après toi, dans cette situation, le petit garçon, il pense que c'est quelque chose qui lui arrive souvent* »).

Enfin, le questionnaire visant à connaître la perception générale de l'agent comporte 16 propositions présentées sous forme d'échelle de Likert en 5 points. Cinq items correspondent aux items « appréciation » du Godspeed, quatre items correspondent aux items « anthropomorphisme » du Godspeed. Un item se rapporte à l'intelligence perçue de l'agent, cinq items visent à mesurer l'apport perçu de l'agent et notamment de son expressivité pour la narration et enfin un item portait sur l'estimation de l'enfant de son propre comportement concernant l'attention portée à l'agent au cours de l'histoire.

### 3.2. Résultats

À partir des mesures décrites ci-dessus, nous discutons de l'impact de l'expression faciale des différentes variables d'*appraisal* manifestée par l'agent sur 1) la mémorisation des événements de l'histoire par les enfants, 2) leur évaluation des différentes séquences émotionnelles et 3) leur appréciation du narrateur virtuel.

Les résultats présentés dans cette section sont statistiquement significatifs lorsque  $p < 0.05$ , et sont considérés comme tendancielles lorsque le  $p$  est compris

---

3.Y compris les dimensions qui n'ont pas été annotées chez l'agent : en effet, c'est l'ensemble de l'expression affective qui est évaluée par l'enfant, indépendamment de notre annotation.

entre 0.05 et 0.1. La petite taille de l'échantillon nous a amené à utiliser le test de Student pour échantillon indépendant afin de traiter les données. Toutes les analyses ont été effectuées à l'aide du logiciel Statistica 9.

En moyenne, les enfants répondent en 23 minutes à l'ensemble des questions et la narration dure une dizaine de minutes. Cette durée d'activité reste raisonnable au regard des capacités attentionnelles des enfants âgés de 6 à 10 ans. Par ailleurs, aucun enfant n'a manifesté de désintérêt pour l'activité.

### 3.2.1. Expressions faciales des différentes variables d'appraisal et mémorisation de l'histoire

Afin d'étudier l'impact des expressions faciales des évaluations, les réponses des 21 questions qui traitent de la mémorisation des événements de l'histoire ont été moyennées pour obtenir une mesure de la mémorisation globale de l'événement de l'histoire pour chaque sujet. Aucun effet significatif n'a été trouvé (voir tableau 1).

Tableau 1. Effet de l'expressivité et de l'âge sur la mémorisation des événements de l'histoire

	T-test	Moyenne	Écart type	Min	Max
Condition « Non-expressive »	P = 0,58	0,65	0,15	0,44	0,88
Condition « expressive »		0,68	0,14	0,46	0,93
Enfants âgés de moins de 8 ans	P < 0,001	0,60	0,12	0,44	0,81
Enfants âgés de plus de 8 ans		0,78	0,09	0,68	0,93

L'âge de nos participants varie entre 6 et 10 ans (14 enfants ont moins de 8 ans). Un T-test a été mené pour explorer l'effet de l'âge sur la mémorisation globale des événements. Une différence significative a été observée entre les deux groupes d'enfants (ceux âgés de moins de 8 ans versus ceux âgés de plus de 8 ans ( $p < 0,001$ )).

Les enfants plus âgés mémorisent mieux les événements de l'histoire que les enfants plus jeunes (voir tableau 1). Ce résultat pourrait expliquer pourquoi aucune différence significative n'a été trouvée entre la mémorisation des événements dans des conditions expressives et non expressives. Compte tenu de la forte influence de l'âge sur la mémorisation des événements de l'histoire, on peut conclure que davantage de participants sur une tranche d'âge mieux ciblée seront nécessaires pour étudier les effets des expressions émotionnelles sur la mémorisation des enfants.

Par ailleurs, quand on compare les deux conditions expressives uniquement pour les enfants les plus jeunes, il apparaît que c'est dans la condition non expressive que les enfants répondent le mieux aux questions qui nécessitent de produire des

inférences quand l'agent virtuel attire leur attention sur des éléments de l'histoire ( $p < 0,05$ ). Il semblerait que les expressions faciales perturbent la compréhension des événements de l'histoire des enfants les plus jeunes lorsque l'agent initie l'attention conjointe.

Cependant, ce résultat ne se retrouve pas pour les enfants les plus âgés pour lesquels aucune différence significative ne s'observe. Néanmoins la taille de l'échantillon reste très limitée, et les résultats descriptifs pour les plus âgés semblent suggérer que l'expressivité faciale favorise la compréhension de l'histoire.

### 3.2.2. Expressions faciales des différentes variables d'*appraisal* et évaluation des événements affectifs réalisée par les enfants

Nous avons appliqué un T-test pour chacune des variables d'*appraisal* (soit 40 réponses : 8 événements \* 5 variables d'évaluation, voir section 3.1) pour étudier l'impact des expressions faciales de l'agent sur l'évaluation des variables d'*appraisal* par les enfants. A nouveau, aucun effet significatif n'a été trouvé.

Un T-test a également été réalisé pour chacun des groupes d'âge pour étudier l'impact de l'expressivité sur l'évaluation des différentes variables d'*appraisal*. Pour les enfants les plus jeunes, la présence d'expressions faciales associées à des variables d'*appraisal* pendant la narration modifie l'évaluation que les enfants réalisent concernant la variable d'*appraisal* « nouveauté » ( $p < 0,05$ ). Dans l'ensemble, les jeunes enfants ont tendance à évaluer différents événements comme plus nouveaux lorsque l'agent virtuel exprime les différentes variables d'*appraisal* pour évaluer un événement émotionnel que quand il ne les exprime pas. Ainsi, la présence d'indices d'*appraisal* via les expressions faciales accentue l'évaluation de la nouveauté. Par exemple, l'événement « *se mettre en rang avant de rentrer en classe* » a été évalué comme prévu dans l'état non expressif (dans un contexte général, il est effectivement attendu de se mettre en rang au moment de la sonnerie) alors qu'il a été évalué comme significativement moins attendu dans l'état expressif (dans le contexte de l'histoire, l'attitude des personnages dans le rang est en effet assez surprenant : ils « font semblant » que tout est normal alors qu'ils leur manque leurs affaires restées bloquées sur le toit).

Un résultat équivalent s'observe concernant la variable « norme ». Dans la condition « expressive », les jeunes enfants ont tendance à considérer quelques événements comme étant moins en conformité avec les normes sociales que dans la condition « non expressive ».

Par exemple dans l'exemple de l'événement « *se mettre en rang avant de rentrer en classe* », dans cette dernière condition, les enfants ont tendance à évaluer cet événement comme étant normal et à ne pas considérer que les protagonistes sont sans chaussure ou sans cartable ce qui est plutôt original.

Enfin, nous avons comparé l'évaluation réalisée par les enfants concernant les variables d'*appraisal* avec l'annotation manuelle des annotateurs. Nous avons constaté que l'évaluation faite par les enfants de ces variables d'évaluation peut varier selon le point de vue considéré. Par exemple, l'événement « l'enseignant

crie » serait surprenant dans une situation générale, mais pourrait être attendu dans le contexte de cette histoire. Ainsi il ressort que l'évaluation des enfants peut être centrée soit sur le contexte spécifique de l'histoire soit être centrée sur un contexte général de la situation (dans l'absolu, tous les matins, les enfants se mettent en rang pour rentrer en classe). Il semblerait que l'expressivité de l'agent oriente le choix du point de vue adopté. Ainsi notre analyse statistique montre la tendance suivante : en condition expressive, les enfants adoptent un point de vue qui se focalise sur le contenu propre de l'histoire pour répondre aux questions alors que dans la condition non expressive, ils adoptent un point de vue plus général. Ces premiers résultats suggèrent que l'expression faciale des variables d'*appraisal* a aidé les enfants à se référer au contenu de l'histoire pour évaluer les événements émotionnels. Lorsqu'aucune expression faciale n'a été affichée, les enfants ont tendance à évaluer les événements selon un point de vue non spécifique à l'histoire. Dans l'ensemble, ces conclusions suggèrent des impacts potentiels de l'affichage des expressions faciales des variables d'*appraisal* sur l'évaluation des événements émotionnels par les enfants.

### 3.2.3. Expressions faciales des différentes variables d'*appraisal* et appréciation de l'agent narrateur par les enfants

Aucune différence significative n'est observée en fonction de l'expressivité de l'agent. Néanmoins, un effet de l'âge s'observe à nouveau. Les enfants les plus jeunes ont tendance à évaluer l'agent narrateur comme étant plus ou moins humain (moyen = 2,6 avec 0 signifiant « machine » et 4 « humain ») alors que les enfants âgés ont tendance à évaluer l'agent comme étant moins humain (moyenne = 1,7). Cependant, les enfants les plus âgés ont tendance à considérer l'agent comme plus sympathique (moyenne = 3,6) que les enfants les plus jeunes (moyen = 3,07). Ce résultat pourrait s'expliquer par le fait que les jeunes enfants ont tendance à percevoir l'agent comme étant plus vivant et ont donc des attentes élevées à son sujet. Les enfants âgés seraient peut-être plus conscients que l'agent était un logiciel informatique et seraient donc plus tolérants vis-à-vis de l'agent narrateur.

## 4. Conclusion et discussion

Notre architecture informatique permet de générer un narrateur virtuel expressif doté d'une expressivité faciale émotionnelle qui caractérise les différentes séquences d'évaluation d'un événement émotionnel selon différents points de vue (point de vue du narrateur, point de vue du narrateur dans un état empathique ou encore point de vue d'un protagoniste de l'histoire). La modélisation puis la génération des comportements expressifs du narrateur virtuel ont été permises en s'appuyant sur le modèle d'*appraisal* de Scherer (2001) qui propose notamment un lien entre les unités d'actions faciales du FACS d'Ekman et Friesen (1975) et les différentes composantes de l'évaluation cognitive des événements émotionnelles. Par ailleurs, le narrateur virtuel est également capable d'initier l'attention conjointe de son auditoire à partir de sa gestualité et de ses expressions.

Nos premiers résultats semblent indiquer que la compréhension de l'histoire ou des événements émotionnels n'est pas affectée par l'expressivité de l'agent. Néanmoins la taille de notre échantillon reste relativement limitée et cela est d'autant plus vrai qu'il y a une grande variabilité en termes d'âge à l'intérieur de l'échantillon. Par ailleurs un effet d'âge s'observe : les enfants les plus jeunes rencontrent plus de difficulté pour comprendre l'histoire que les enfants les plus âgés. L'expressivité de l'agent pour les enfants les plus jeunes semble les perturber quand il s'agit d'élaborer une représentation globale de l'histoire mais semble les aider à choisir un point de vue plus cohérent quand il s'agit d'évaluer des événements émotionnels. En effet, ils adoptent moins face à un agent expressif un point de vue général sur les événements de l'histoire qui s'appuierait sur leurs éléments de vie personnelle mais privilégient un point de vue spécifique centré sur le contenu de l'histoire. Concernant les enfants les plus âgés, les résultats ne sont pas significatifs et la taille de l'échantillon est très limitée 3 ou 4 enfants par condition. Néanmoins, les résultats descriptifs suggèrent que l'expressivité soutiendrait plutôt la compréhension de l'histoire. Ainsi ces premiers résultats nous invitent à poursuivre ce travail en étendant la taille de notre échantillon d'une part, et d'autre part, investiguer une approche développementale pour étudier dans quelle mesure l'expression faciale des variables d'*appraisal* de la part du narrateur peut contribuer à une meilleure compréhension des récits.

Dans le présent travail, l'évaluation des variables d'*appraisal* et les différents points de vue à adopter ont été réalisés à partir d'une annotation manuelle. Ce processus est assez coûteux (une journée par personne pour annoter une histoire d'une centaine de phrases) et une évaluation automatique des variables d'*appraisal* représenterait une étape importante pour la conception et la génération de narrateurs virtuels entièrement automatiques. En effet, une évaluation automatique de ces composantes en fonction du contenu de l'histoire et de la mémorisation de l'événement par l'agent pourrait être d'un grand intérêt pour le développement d'un modèle informatisé de narration expressive autonome.

Une autre limite de notre cadre interactif actuel est qu'il est basé sur une approche « Wizard of Oz ». Le narrateur virtuel peut écouter la réponse de l'enfant après des énoncés interrogatifs. Cependant, le traitement des interruptions réalisées par l'enfant pendant la tâche de récit augmentera la qualité de l'interaction.

Outre la taille de l'échantillon, il est possible que l'absence de résultat relève d'un manque d'intérêt porté par les enfants au narrateur virtuel. Afin d'étudier cette question, nous utiliserons dans la prochaine étude un système d'oculométrie afin de renseigner la prise d'informations réalisée par les enfants au cours de l'interaction narrative. Cela permettra d'enregistrer et d'étudier la fréquence et la durée des regards que les enfants consacrent au visage de l'agent narrateur vs au graphisme de l'histoire et de voir si cela varie en fonction de l'expressivité émotionnelle de l'agent.

Enfin, dans un travail futur à plus long terme, nous visons également à enrichir la communication non verbale affective du narrateur virtuel. En effet, le narrateur virtuel actuel est doté d'une voix humaine exprimant des émotions. Il s'agit donc



d'une limite du travail actuel puisque cela peut conduire à une incohérence perceptive lorsque l'agent n'exprime pas d'émotions sur son visage mais en exprime vocalement. Cet enrichissement de la communication non verbale devra donc notamment porter sur les aspects prosodiques mais également sur l'expression corporelle. Cela permettra d'une part de soutenir les spécifications pour la conception d'agent virtuel narratif plus crédible et d'autre part de mieux comprendre comment une approche multimodale de la narration peut étayer la compréhension des récits de la part des enfants et l'élaboration de connaissances plus spécifiques notamment celles portant sur les émotions.

#### *Remerciements*

*Cette recherche a bénéficié du soutien du programme de l'Agence Nationale de la Recherche (ANR), projet NARECA (ANR-13-CORD-0015).*

*Les auteurs tiennent à remercier les enfants et leurs parents pour avoir accepté de participer à cette étude ainsi que le directeur du centre de loisirs et les animateurs pour leur accueil et leur contribution au bon déroulement de l'étude. Enfin, les auteurs souhaitent remercier Matthieu Courgeon pour l'aide apportée dans le développement du conteur virtuel.*

#### **Bibliographie**

- Adam C., Cavedon L. (2013). Once upon a time... a companion robot that can tell stories. Poster paper. In *Intelligent Virtual Agents, Thirteenth International Conference, IVA*, Edinburgh, UK.
- Atkinson RK. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, vol. 94, n° 2, 416-427. doi:10.1037/0022-0663.94.2.416
- Bartneck C., Croft E., Kulic D., Zoghbi S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, vol. 1, n° 1, 71-81. doi: 10.1007/s12369-008-0001-3
- Bickmore T., Trinh H., Hoppmann M., Asadi R. (2016). Virtual Agents in the Classroom: Experience Fielding a Co-Presenter Agent in University Courses, In *Sixteenth International Conference on Intelligent Virtual Agents, IVA*, Los Angeles, California.
- Bleackley P., Hyniewska S., Niewiadomski R., Pelachaud, C. Price M., (2010) Emotional Interaction Storyteller System, *KEER International Conference on KANSEI Engineering and Emotion Research*, Paris, 2-4 March 2010, pp. 1716-1726.
- Bosse T., Pontier M., Siddiqui G. F., Treur J., (2007). Incorporating emotion regulation into virtual stories. In *Intelligent Virtual Agents, 7th International Conference, IVA 2007*, Paris, France, September 17-19, 2007, Proceedings, p. 339-347, 2007.
- Bruner J. (1990). Car la culture donne forme à l'esprit. De la révolution cognitive à la révolution culturelle, Paris, ESHEL.
- Bruner J. (1996). L'éducation, entrée dans la culture. Paris, Retz

- Bruner J. (2000). *Culture et modes de pensée*. Paris : Retz
- Bruner J. (1983). *Le développement de l'enfant : Savoir faire savoir dire*. Paris, PUF.
- Buisine S., Wang Y., Grynszpan O., (2009). Empirical investigation of the temporal relations between speech and facial expressions of emotion. *Journal on Multimodal User Interfaces*, vol. 3, n° 4, p. 263–270.
- Callaway C. B., Lester J. C. (2002). Narrative prose generation. *Artificial Intelligence*, vol. 139, n° 2, p. 213-252.
- Carrera-Levillain P., Fernandez-Dols J. M., (1994). Neutral faces in context: Their emotional meaning and their function. *Journal of Nonverbal Behavior*, vol. 18, n°4, p. 281–299.
- Cavazza M., Charles F., Mead S. J. (2002). Character-based interactive storytelling. In *IEEE Intelligent Systems*, vol. 17.
- Chella A., Barone R.E., Pilato G., Sorbello R. (2008). An Emotional Storyteller Robot. *AAAI Spring Symposium: Emotion, Personality, and Social Behavior 2008*, 17-22
- Courgeon M., Clavel C., Martin JC. (2014). Modeling Facial Signs of Appraisal During Interaction: Impact on Users' Perception and Behavior. *AAMAS* vol. 59, n° 1, 765-772
- Courgeon M., Martin J.C., Jacquemin C. (2008). Marc: a multimodal affective and reactive character. In: *Proceedings of the 1st Workshop on Affective Interaction in Natural Environments*
- Crahay M. (1999). *Psychologie de l'éducation*. Paris, Presses universitaires de France.
- De Jong M.T., Bus A.G. (2004). The efficacy of electronic books in fostering kindergarten children's emergent story understanding. *Reading Research Quarterly*, vol. 39, p. 378-393.
- Doty D.E., Popplewell S.R., Byers G.O. (2001). Interactive CD-ROM storybooks and young readers' reading comprehension. *Journal of Research on Computing in Education*, vol. 33, p. 374-384.
- Ekman P, Friesen W. (1975). *Unmasking the Face: A guide to recognizing emotions from facial clues*. Prentice-Hall, Inc.
- Fridin M. (2014). Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers & Education*, vol. 70, p. 53-64
- Frijda N. (1993). The place of appraisal in emotion. *Cognition and Emotion*, vol. 7, n° 3-4, 357-387;
- Guo Y.R, Hoe-Lian Goh D., Luyt B., Sin SC., Ang R.P. (2015). The effectiveness and acceptance of an affective information literacy tutorial, *Computers & Education*, vol. 87, p. 368-384.
- Kory J., Breazeal C. (2014). Storytelling with Robots: Learning Companions for Preschool Children's Language Development. In P. A. Vargas & R. Aylett (Eds.), *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE: Washington.
- Lester JC., Converse SA., Kahler SE., Todd Barlow S., Stone BA., Bhogal RS. (1997). The persona effect: affective impact of animated pedagogical agents, *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, p.359-366, March 22-27, 1997, Atlanta, Georgia, USA.

- Marsella S.C., Gratch J., Petta P. (2010). Computational Models of Emotion. In: Scherer, K.R., Banziger, T., Roesch (eds.) *A blueprint for affective computing: A sourcebook and manual*, p. 21–46. Oxford University Press, Oxford
- Marsella S.C., Gratch J. (2009). EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, vol. 10, n° 1, p. 70-90.
- Martin J.C., Deviller L., Raouzaïou A., Caridakis G. (2011). Coordinating the generation of signs in multiple modalities in an affective agent. *Emotion-Oriented Systems: The Humaine Handbook, Cognitive Technologies*, p. 349-367.
- Miletitch, R., Sabouret N., Ochs M. (2012). Susciter l'émotion dans la narration automatique, In *Technique et Science de l'Informatique (TSI)*, vol. 31, n° 4.
- Moody A.K., Justice L.M., Cabell S.Q. (2010). Electronic versus traditional storybooks: relative influence on preschool children's engagement and communication. *Journal of Early Childhood Literacy*, vol. 10, n° 3, p. 294-313.
- Mutlu B., Forlizzi J., Hodgins, J. (2006). A storytelling robot: Modeling and evaluation of human-like gaze behavior. In G. Sandini & A. Billard (Eds.), *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots*, p. 518-523.
- Noorhidawati A., Ghazal Ghalebani S., Siti Hajar R. (2015). How do young children engage with mobile apps? Cognitive, psychomotor, and affective perspective, *Computers & Education archive*, vol. 87, n°C, p. 385-395.
- Oatley K. (2004). Scripts, Transformations, and Suggestiveness of Emotions in Shakespeare and Chekhov. *Review of General Psychology*, vol. 8, n°4, p. 323-340.
- Oberlander J., Robertson J. (2002). Ghostwriter: Educational Drama and Presence in a Virtual Environment. *Journal Computer-Mediated Communication*, vol. 8.
- Ochs M., Niewiadomski R., Pelachaud C. (2010). How a virtual agent should smile? Morphological and dynamic characteristics of virtual agent's smiles. In *Intelligent Virtual Agents*, 10th International Conference, IVA, p. 427-440, Philadelphia, Pennsylvania, USA.
- Raimundo G., Cabral J., Melo C., Oliveira L.C., Paiva A. (2007). Telling Stories with a Synthetic Character: Understanding Inter-modalities Relations, *COST workshop, Verbal and Nonverbal Communication Behaviours Lecture Notes in Computer Science*, vol. 4775/2007.
- Ricci C.M., Beal C.R. (2002). The effect of interactive media on children's story memory. *Journal of Educational Psychology*, vol. 94, n° 1, p. 138-144.
- Riedl M. O. (2004). *Narrative generation: balancing plot and character*. North Carolina State University.
- Rimé B. (1997). Cognition et émotion in J.-P. Leyens & J.-L. Beauvois (Eds.), *La psychologie sociale 3 : L'ère de la cognition Grenoble*, p. 107-125, Presses Universitaires de Grenoble.
- Scherer K. (2001). Appraisal considered as a process of multilevel sequential checking. Dans K. Scherer, A. Schorr, & T. Johnstone, *Appraisal processes in emotion* (pp. 92-121). New York, Oksford University Press.
- Serban O., Pauchet A., Bersoult A., Chanoni E., (2013). Tell Me a Story : A Comparative Study of Child Interaction with Virtual Characters and Adults. R. Aylett et al. (Eds.), *13th*

*international conference on Intelligent Virtual Agents (IVA)*, LNAI 8108, pp. 459-460, Springer, Edinburgh, UK, 2013.

Silva A., Raimundo G., Paiva A. (2003). Tell me that Bit Again... Bringing Interactivity to a Virtual Storyteller, *Proceedings of the 2nd International Conference on Virtual Storytelling - ICVS 2003*.

Silva A., Vala M., Paiva A. (2001). Papous: The Virtual Storyteller In Intelligent Virtual Agents, *International Conference IVA 2001*, Springer.

Swartjes I. M. T., Theune M. (2008). The virtual storyteller: Story generation by simulation.

Szafir D., Mutlu B. (2012). Pay attention!: designing adaptive agents that monitor and improve user engagement, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, May 05-10, 2012, Austin, Texas, USA.

Theune M., Linssen J., Alofs T. (2013). Acting, Playing, or Talking about the Story : An Annotation Scheme for Communication during Interactive Digital Storytelling. *6th International Conference on Interactive Digital Storytelling*, p. 132-143.

Vygotski L. (1997). *Pensée et Langage*. Paris, La Dispute.