

INTERAZIONE AFFETTIVA CON AGENTI ANIMATI

Fiorella De ROSIS, Addolorata CAVALLUZZI, Berardina De CAROLIS

Dipartimento di Informatica — Università di Bari
Via Orabona, 4
70125 Bari - Italy
Mél : {derosis, cavalluzzi, decarolis}@di.uniba.it

Riassunto

Partendo dal presupposto che l'interazione con la tecnologia sia guidata da regole mutuabili dalla teoria delle relazioni sociali, ci siamo poste l'obiettivo di verificare se l'utilizzo di nuove metafore di interazione, come quella degli agenti animati conversazionali, contribuisca a rafforzare questo effetto. In questo lavoro, descriviamo i risultati di una serie di studi di valutazione realizzati allo scopo di identificare le reazioni degli utenti a messaggi persuasivi 'one-shot' o a dialoghi con un agente animato in cui veniva esaminato, in particolare, il ruolo giocato dall'espressione, sia verbale che non verbale, delle emozioni. Gli studi ci hanno dato indicazioni interessanti sul ruolo dell'espressione delle emozioni ma ci hanno anche permesso di concludere che la strada da percorrere per la realizzazione di nuove metafore d'interazione in cui la tecnologia non giochi il ruolo di 'strumento' ma quello di 'assistente amichevole o empatico' sia ancora lunga da percorrere.

Abstract

AFFECTIVE INTERACTIONS WITH EMBODIED AGENTS

Starting from the postulate according to which interaction with technology depends on rules specific to the theory of social relations, our goal was to examine whether the use of new metaphors of human-machine interaction, like that given by animated conversational agents, can contribute to reinforcing this effect. In this work, we describe the results of a series of evaluation studies performed in order to identify users' reactions to persuasive "one-shot" messages or to dialogues with an animated agent. In particular, the goal was to examine within this framework the role played by the expression (both verbal and non-verbal) of emotions. These studies have provided us with interesting information on the role of the expression of emotions and have also allowed us to conclude that the path to follow in developing new interaction metaphors appropriate for conceiving of the machine not only as an "instrument," but also as a "friendly or empathetic assistant" is still a long way off.

Resumo

INTERAÇÕES AFETIVAS COM AGENTES ANIMADOS

Começando pelo postulado segundo o qual a interação com a tecnologia depende de regras específicas à teoria de relações sociais, nosso objetivo foi examinar se o uso de novas metáforas de interação homem-máquina, como aquela dada por agentes conversacionais animados, pode contribuir para reforçar esse efeito. Neste trabalho, descrevemos os resultados de uma série de estudos de avaliação conduzidos para identificar a reação dos

usuários com mensagens do tipo "one-shot" ou com diálogos com um agente animado. Mais especificamente, o objetivo foi examinar dentro dessa abordagem o papel desempenhado pela expressão de emoções (tanto verbais quanto não-verbais). Estes estudos nos deram informação pertinente quanto ao papel da expressão das emoções e também nos permitiu concluir que ainda é longo o caminho a seguir para desenvolver novas metáforas interacionais apropriadas para conceber a máquina não apenas como instrumento, mas também como assistente pelo qual teríamos algum tipo de empatia ou relação amical.

Resumen

INTERACCIONES AFECTIVAS CON AGENTES ANIMADOS

Basándonos en el postulado que la interacción con la máquina depende de reglas propias a la teoría de las relaciones sociales, nos fijamos como objetivo comprobar si la utilización de nuevas metáforas de la interacción persona-máquina, como aquella proporcionada por los agentes conversacionales animados, podía contribuir a fortalecer este efecto. En este trabajo, describimos los resultados de una serie de trabajos de evaluación realizados con el objetivo de identificar las reacciones de los usuarios a mensajes persuasivos one-shot o a diálogos con un agente animado. Más particularmente, se trató de examinar en este marco el papel desarrollado por la expresión de emociones, sea verbal o no verbal. Estos estudios nos proporcionaron indicaciones interesantes sobre el papel de la expresión de las emociones pero también nos permitieron concluir que queda mucho camino por recorrer hacia la realización de nuevas metáforas de interacción aptas a concebir la máquina no sólo como 'instrumento' pero como 'asistente amical o empático'.

Résumé

INTERACTIONS AFFECTIVES CHEZ LES AGENTS ANIMÉS

Partant du postulat selon lequel l'interaction avec la machine s'appuie sur des règles propres à la théorie des relations sociales, nous nous sommes données comme objectif de vérifier si l'utilisation de nouvelles métaphores de l'interaction personne-machine, comme celle donnée par les agents conversationnels animés, pouvait contribuer à renforcer cet effet. Dans ce travail, nous décrivons les résultats d'une série de travaux d'évaluation réalisés dans le but d'identifier les réactions des utilisateurs à des messages persuasifs one-shot ou à des dialogues avec un agent animé. Plus particulièrement, il s'agissait d'examiner dans ce cadre le rôle joué par l'expression des émotions, qu'elle soit verbale ou non verbale. Ces études nous ont fourni des indications intéressantes sur le rôle de l'expression des émotions, mais elles nous ont également permis de conclure que le chemin à parcourir vers la réalisation de nouvelles métaphores d'interaction aptes à concevoir la machine non plus seulement comme « instrument » mais comme « assistant amical ou empathique » était encore bien long.

1. Introduzione

Gli agenti animati sono personaggi virtuali che possono assumere l'apparenza di cartoni animati o di personaggi più o meno realistici, rappresentati mediante immagini bi o tridimensionali. Qualunque sia il loro aspetto, questi personaggi sono dotati di personalità e della capacità di mostrare, nel viso, nei gesti e nella voce, espressioni simili a quelle che gli umani mostrano nei loro processi di comunicazione (Cassell et al., 2000). Possono essere dotati di proprietà varie, a seconda del dominio applicativo: dalla pronuncia di monologhi alla capacità di dialogare con l'utente (nel

qual caso, si parla di 'agenti conversazionali'). Quella con gli agenti animati è vista come una nuova metafora d'interazione con i mezzi di calcolo, che sostituisce l'idea di 'usare uno strumento' (tipica delle interfacce grafiche o a manipolazione diretta) con quella di 'interagire con un amico o un assistente' capace di sostenere l'utente nello svolgimento dei suoi compiti e di aiutarlo in caso di difficoltà. Nel paradigma 'Computers As Social Actors' (CASA), sostenuto da un gruppo di ricercatori dell'Università di Stanford, si afferma che "one can take both theories and methods from social psychology and directly apply them to human-technology interaction": cioè che l'interazione con la tecnologia sarebbe guidata da regole mutuabili dalla teoria delle relazioni sociali. Questo paradigma si basa sui risultati di una serie di esperimenti orientati a studiare le risposte di soggetti umani a forme d'interazione con la tecnologia basate sia su interfacce in linguaggio naturale che su agenti animati. Gli esperimenti studiavano l'influenza, sull'efficacia dell'interazione, di fattori quali la 'somiglianza', tra utenti e interfaccia, nel genere nell'etnia e nella personalità e anche l'effetto del simulare, nell'interfaccia, elementi di adulazione o di gentilezza (Nass *et al.*, 1995; Nass *et al.*, 1997; Fogg e Nass, 1997). La regola di applicabilità delle teorie sociali all'interazione con la tecnologia sembra essere valida anche quando vengono impiegati i metodi d'interazione oggi più comuni, come la 'manipolazione diretta' di oggetti grafici. Ma aumentano, evidentemente, quando le forme di interazione si basano su un'antropomorfizzazione della tecnologia stessa: cioè, quando il calcolatore utilizza forme di comunicazione che sono tipiche degli umani: una faccia o un corpo con sembianze più o meno naturali, una voce, il linguaggio naturale. L'affettività è un elemento essenziale delle relazioni sociali. Può coinvolgere aspetti relativamente stabili, come la personalità, o più dinamici, come l'emotività o l'umore; può riguardare sensazioni individuali, come l'ansia, la frustrazione o la soddisfazione, o sociali, come l'empatia, la gelosia, l'ammirazione (Ortony *et al.*, 1988). Ha ricevuto, in passato, un'attenzione limitata nell'ambito dell'intelligenza artificiale ma è l'oggetto, negli ultimi anni, di un interesse crescente: vedi ad esempio (Picard, 2000, 2002; De Rosis, 2001, 2002) e il Network Europeo di Eccellenza sulle Emozioni, HUMAINE¹.

All'interno dei possibili ambiti di applicazione del paradigma CASA, abbiamo considerato, in particolare, il problema della simulazione di processi di persuasione orientati a convincere l'interlocutore a compiere specifiche azioni. La simulazione di questi processi ha un ampio spettro di teorie cognitive e psico-linguistiche a cui far riferimento: ad esempio, gli schemi argomentativi di (Walton, 1996a,b) che descrivono i diversi modi in cui si può argomentare, e la teoria degli *stage of change* (Prochaska *et al.*, 1992) che descrive quali aspetti dello stato cognitivo dell'interlocutore influenzino l'efficacia delle diverse strategie argomentative. Nel nostro generatore di messaggi persuasivi, i testi combinano elementi razionali con elementi emotivi (Carofiglio e De Rosis, 2003; Miceli *et al.*, in press), formulano il messaggio in modo negativo o positivo a seconda del contesto e possono essere realizzati in linguaggio naturale oppure essere pronunciati da un agente animato, mentre i dialoghi pianificano contenuto e forma della 'prossima mossa' in funzione di ipotesi sullo stato cognitivo ed emotivo dell'interlocutore (Cavalluzzi *et al.*, 2004). In questo lavoro, descriviamo i risultati di una serie di studi di valutazione realizzati allo scopo di identificare le reazioni degli utenti a messaggi persuasivi *one-shot* pronunciati dall'agente, o a dialoghi con un agente animato in cui viene esaminato, in

¹ <http://emotion-research.net>

particolare, il ruolo giocato dall'espressione, sia verbale che non verbale, delle emozioni.

2. Due studi sperimentali

Ci siamo posti il quesito seguente: *è ragionevole assumere che i fattori affettivi giochino, nell'interazione con la tecnologia, un ruolo simile a quello giocato nei rapporti fra gli esseri umani, almeno limitatamente ad alcuni settori, come quello dell'informazione e della persuasione? E l'utilizzo di nuove metafore dell'interazione, come quella degli agenti animati conversazionali, contribuisce a rafforzare questo effetto?* Per tentare di dare una risposta (anche se parziale) a questo quesito, abbiamo realizzato due studi, che si ponevano obiettivi diversi:

1. in un primo studio, basato su monologhi, abbiamo studiato gli effetti sull'utente di variazioni nel contenuto emotivo dei messaggi e nelle loro modalità di espressione, misurati in termini di informatività, credibilità e persuasività percepita del messaggio;
2. in un secondo studio, basato su dialoghi, abbiamo studiato la relazione sociale che si stabilisce fra utente ed agente animato, e come questa dipenda dalle caratteristiche dell'utente e (di nuovo) dalle modalità di espressione di cui l'agente è dotato.

Il dominio applicativo è quello dei suggerimenti sull'alimentazione. Nello studio sui monologhi, il messaggio era orientato, in particolare, a persuadere l'interlocutore a seguire una dieta che contenga una giusta dose giornaliera di vegetali, evidenziandone i vantaggi e spiegandone le ragioni. Nello studio sui dialoghi, l'obiettivo era esteso alla descrizione dei componenti principali della dieta (carboidrati, proteine ecc) e al mettere in guardia l'interlocutore nei confronti di comportamenti alimentari poco corretti.

3. Studio su monologhi

Il primo studio di valutazione è stato effettuato nel corso del Progetto Europeo Magicster (Berry *et al.*, in press). Aveva come ipotesi che l'effetto persuasivo di un messaggio sia influenzato da almeno due fattori: da un lato, l'attivazione nell'interlocutore di reazioni emotive sia di tipo positivo che di tipo negativo, e dall'altro l'espressione, da parte di chi deve persuadere, di emozioni coerenti con il contenuto del messaggio persuasivo.

3.1 Metodo

Come suggerito da (Nass *et al.*, 2000), due principi guidano il disegno degli studi nella ricerca sperimentale: *varied conditions* e *random assignment*. Il principio di *varied conditions* suggerisce di non limitarsi a mostrare ai soggetti coinvolti nello studio un prototipo sottoponendo loro un questionario di valutazione. Afferma che occorre, piuttosto, fare delle ipotesi precise sui fattori che possono condizionare la percezione del prototipo, realizzare prototipi diversi che corrispondano a ciascuna di esse, sottoporre ogni prototipo ad un diverso gruppo di soggetti e quindi confrontare le valutazioni effettuate dai diversi gruppi. *Random assignment* denota il principio classico secondo cui l'assegnazione dei soggetti coinvolti nello studio ad ogni gruppo è effettuata in modo casuale. Lo studio che abbiamo realizzato aveva, come disegno,

quello dei cosiddetti 'studi comparativi' di tipo *between subject* e ha coinvolto, complessivamente, 178 soggetti. In ogni fase, siamo partiti da un'ipotesi, abbiamo formulato due o più messaggi da mettere a confronto, formato due o più gruppi di soggetti ai quali mostrare ciascuno dei messaggi e valutato con un questionario i risultati prodotti dal messaggio in ciascuno dei soggetti. L'obiettivo persuasivo era quello di 'convincere il soggetto a mangiare vegetali'. Il messaggio era persuasivo nel senso che si prospettavano quali fossero i vantaggi di questa dieta (quando il messaggio era formulato in modo positivo) o gli svantaggi di non seguire la dieta (quando il messaggio era formulato in modo negativo). Quindi il messaggio cominciava con una descrizione della fonte, delle dosi suggerite e del significato di 'dose':

"The Food Standards Agency in the UK, and Departments of Health throughout the Western World, recommend that people should aim to eat at least five portions of fruit and vegetables a day. These can be fresh, frozen, tinned, dried or juiced, although eating them raw or lightly cooked is best. Potatoes should not be counted because they contain starch. A recommended portion is roughly 80g, or for example, an apple, 2-3 tablespoons of frozen vegetables, or 1 glass of fruit juice. Children can be given smaller portions, according to their appetite. Although the recommended amounts may seem a lot, fruits and vegetables are good sources of vitamins (particularly A and C), minerals, and fibre."

e continuava poi con la parte più specificamente persuasiva: effetti dei costituenti dei vegetali (vitamina A e C, minerali e fibre) sulla salute e sull'aspetto fisico:

"As far as vitamins are concerned, research has shown that eating the recommended levels of vitamin A and C can have beneficial effects for you appearance and health. These vitamins not only cleanse the blood but are important for growth and the repair of body tissues. In terms of your appearance, eating foods which are rich in vitamin A and C can lead to good skin and hair, as vitamin C stimulates, tonifies and rehydrates them. Eating these vitamins will also result in healthy gums and strong teeth. Diets that are rich in vitamin A and C are also good for your health. They give resistance against colds, flu and other infections, and generally result in higher energy levels. They are also associated with a healthy heart and stomach. Consuming extra amounts of these vitamins has been shown to relieve asthma, improve blood sugar control, and promote healing. Vitamin A is essential for good vision, particularly in dim lighting conditions. Many fruit and vegetables also contain minerals such as calcium, which are important for strong bones and teeth, effective muscles, and a well-functioning nervous system. They are also a good source of fibre. Eating sufficient quantities of fibre is important for a healthy heart and effective digestive system, and can improve dental well-being."

Abbiamo effettuato tre tipi di confronti:

3.1.1. Framing negativo vs positivo

L'idea era, in questo caso, verificare quanto valga, nel dominio della persuasione a seguire una dieta 'corretta', il cosiddetto '*framing effect*', e in particolare la valenza con la quale il messaggio persuasivo viene formulato. (Levin *et al.*, 1998) distinguono tre tipi di *framing*:

- il cosiddetto *risky choice*, introdotto originalmente da (Tversky e Kahneman, 1981), che coinvolge la descrizione degli "outcomes of a potential choice involving options differing in level of risk". Ad esempio: "asure saving of one-third of the lives" vs "asure loss of two-thirds of the lives";
- l'*attribute framing*, che riguarda la descrizione delle caratteristiche di oggetti o eventi. Ad esempio, "il maiale contiene l'85 % di carne magra" oppure "il 15% di grassi";
- il *goal framing*, che coinvolge la descrizione degli obiettivi per i quali una certa azione suggerita dovrebbe essere eseguita.

Nei tre casi, la scelta di rischio, gli attributi dell'oggetto o il *goal* da raggiungere vengono descritti, nel *positive* e nel *negative framing*, rispettivamente in termini positivi o negativi. Nei testi preparati per questo studio erano inseriti diversi elementi di *goal framing*. Si faceva cioè leva, nel *positive framing*, sull'obiettivo di 'realizzare conseguenze positive' eseguendo l'azione suggerita. Nell'esempio del paragrafo precedente: "As far as vitamins are concerned, research has shown that eating the recommended levels of vitamin A and C can have beneficial effects for you appearance and health. These vitamins not only cleanse the blood but are important for growth and the repair of body tissues". Nel *negative framing*, invece, si faceva leva sull'obiettivo di 'evitare conseguenze negative' del non compierla. Ad esempio: "As far as vitamins are concerned, research has shown that failure to eat the recommended levels of vitamin A and C can have detrimental effects for you appearance and health. A shortage of these vitamins will impede blood cleansing and impair growth and the repair of body tissues." Secondo lo schema persuasivo che Walton chiama 'from consequences' (Walton, 1996a,b), il messaggio conteneva, nel primo caso, l'idea di far ricorso alla speranza di star meglio in salute ed avere un aspetto migliore, mentre nel secondo caso si faceva ricorso alla paura, mostrando gli effetti del non compiere l'azione. C'erano, inoltre, nel messaggio diversi elementi di *attribute framing*, consistenti con la scelta del *goal framing*. I due messaggi erano formulati in modo tale che il loro contenuto e la loro lunghezza fossero identici e che fossero diversi solo gli aggettivi usati e la formulazione della frase. Questa fase dello studio ha coinvolto, complessivamente, 46 soggetti (due gruppi di 23 soggetti ciascuno) ai quali abbiamo mostrato un messaggio in linguaggio naturale sul video del computer: ad un gruppo veniva mostrato il messaggio positivo e all'altro quello negativo. Dopo aver letto il messaggio, ad ogni soggetto veniva chiesto di compilare un questionario che conteneva un elenco di domande relative al messaggio indicando, per ogni domanda, un valore in una scala di Likert (da 1 a 6):

- On the basis of what you have just read how likely are you to follow the recommendations for healthy eating in the future?
- How satisfied were you with the information contained in the message?
- How easy did you think that the message was to understand?
- How persuasive did you think the healthy eating message was?
- How credible did you think the message was?
- To what extent do you feel that there is good evidence to support the recommendations in the message?
- Overall, how would you rate the quality of the arguments contained in this message?
- How trustworthy did you think the content of the message was?
- How sincere did you think the source of the message was?

3.1.2. Contenuto e stile razionale vs emotivo

In questa fase, l'obiettivo era confrontare come la qualità e l'impatto persuasivo di un messaggio siano influenzati dal suo contenuto apertamente emotivo. Anche in questo campo, diversi studi dimostrano le differenze di risposta a messaggi 'razionali' e 'emotivi' in diverse condizioni di motivazione e di capacità di elaborare il messaggio (Petty e Cacioppo, 1986; Rosselli *et al.*, 1995). Uno dei punti ancora da approfondire riguarda i processi di elaborazione dell'informazione ricevuta, nei due casi: presumibilmente, elaborazioni 'cognitive' nei messaggi razionali, e una combinazione di elaborazioni cognitive ed affettive nei messaggi emotivi (ancora Rosselli *et al.*, 1995). Nello studio sono stati coinvolti 60 soggetti, divisi in 2 gruppi da 30. Il messaggio positivo precedente (che noi consideravamo essere sostanzialmente 'razionale'), era messo a confronto con un messaggio anch'esso positivo, cioè orientato a mostrare gli effetti positivi del mangiare vegetali, ma in cui erano inseriti 'contenuti emotivi'. La premessa di ogni parte persuasiva era identica, nei due messaggi: "*As far as vitamins are concerned, research has shown that eating the recommended levels of vitamin A and C can have beneficial effects for you appearance and health. These vitamins not only cleanse the blood but are important for growth and the repair of body tissues*". Nel seguito, orientato a 'giustificare' le ragioni dell'affermazione precedente, venivano invece inserite delle differenze. Nel messaggio 'razionale', sostanzialmente si spiegavano le cause di ciascuno degli effetti positivi evidenziati: nell'esempio, la vitamina A fa bene alla pelle perché fa bene alla differenziazione cellulare. Nel messaggio 'emotivo', invece, l'obiettivo era stimolare una visualizzazione degli effetti. La differenza fra i due messaggi era ulteriormente rafforzata dall'uso di termini 'pseudo-scientifici' nel primo caso e di aggettivi 'attraenti' nel secondo. Questa formulazione dei messaggi emotivi riproduceva l'idea che "*Apersuasive message might be considered as emotional if it served to cue the retrieval of emotional experiences associated with an object. In other words, emotional appeals can be thought of as attempts to elicit, or make salient, object-related emotions*" (Cohen, 1990).

Esempio:

[RATIONAL] In terms of your appearance, eating foods which are rich in vitamin A and C can lead to strong and healthy hair, as vitamin C neutralizes substances which damage them. By reducing cell damage, these vitamins will also result in healthy gums and strong teeth that look and feel noticeably better too.

[EMOTIONAL] In terms of your appearance, eating foods which are rich in vitamin A and C can lead to strong and shiny hair, reducing the need for artificial, glycerin-based conditioners. Eating these vitamins will also result in healthy gums and strong teeth, which means that you will be able to laugh and smile with total confidence.

Come si nota, nel messaggio razionale si spiega perché la vitamina C ha effetti positivi sull'aspetto fisico (neutralizza sostanze che danneggiano i capelli), mentre nel secondo se ne visualizzano gli effetti: "pensa che non avrai bisogno di sostanze speciali per rendere i tuoi capelli più belli e più brillanti".

3.1.3. Messaggio pronunciato da un agente animato

In questa fase, ci proponevamo di confrontare l'effetto di un messaggio scritto con quello di un messaggio pronunciato da un agente animato e di studiare quanto

l'espressione, nel viso, di emozioni che coerenti con il contenuto del messaggio, influenzi l'impatto del messaggio stesso. Il messaggio presentato ai soggetti era quello 'razionale e positivo' utilizzato nei due studi precedenti, ed era pronunciato da Greta (figura 1), un agente animato tridimensionale molto realistico che è in grado di mostrare, col viso, espressioni di diverso tipo (*turn taking*, performativi, *belief-relation*, adjectival expressions e emozioni: Poggi e Pelachaud, 2000).



Figura 1 — L'agente animato 'GRETA' di Catherine Pelachaud (Pelachaud e Bilvi, 2003).

Sono stati coinvolti in questa fase tre gruppi di soggetti, di 24 individui ciascuno (72 soggetti in totale). L'agente aveva caratteristiche diverse nei tre gruppi:

- nel primo gruppo, Greta accompagnava il messaggio con le espressioni del viso che sono necessarie per rendere un messaggio coerentemente 'espressivo' con il suo contenuto. L'agente non mostrava però nessuna emozione;
- nel secondo gruppo, Greta mostrava emozioni coerenti con il contenuto del messaggio: quindi, in sostanza accompagnava le parti 'positive' con un sorriso;
- nel terzo gruppo, Greta mostrava emozioni inconsistenti con il contenuto del messaggio: quindi, accompagnava le parti 'positive' del messaggio con un'espressione di tristezza.

In questo studio, oltre a chiedere ai soggetti di rispondere alle stesse domande che erano inserite nel questionario utilizzato nei due studi precedenti, chiedevamo anche una valutazione dell'agente (sempre con una scala di Likert, da 1 a 6):

For each word below, please indicate how well it describes the computer character that you have just seen. (Note that you are now evaluating the character, NOT the message):

Intelligent, Competent, Credible, Reliable, Helpful, Likable

3.2. Risultati

Per quanto riguarda il confronto fra framing negativo e positivo del messaggio, non sembrano esserci differenze significative fra le due modalità per la maggior parte dei fattori esaminati (tabella 1). Il messaggio negativo sembra essere leggermente più credibile, e i soggetti dicono che è un po' più probabile che ne seguiranno i suggerimenti, mentre la qualità degli argomenti usati è ritenuta leggermente migliore per il messaggio positivo, che è considerato più degno di fiducia. Quindi, in sostanza,

la qualità degli argomenti sembra essere maggiore nel messaggio positivo. Ma, nello stesso tempo, il messaggio negativo sembra più efficace nel promuovere il comportamento desiderato. Questi risultati sono coerenti con quelli prodotti da studi precedenti. E' noto infatti, che un *attribute framing* positivo produce 'valutazioni più favorevoli' rispetto ad uno negativo, probabilmente perché promuove l'attenzione selettiva nei confronti di aspetti positivi dell'oggetto o dell'evento descritto la quale produce, a sua volta, l'accesso ad associazioni positive (Levin *et al.*, 1998). Un effetto inverso è stato osservato, invece, nel *goal framing*: in questo caso, diversi studi (vedi ad esempio, Meyerowitz e Chaiken, 1987) hanno dimostrato che i messaggi negativi hanno un impatto persuasivo maggiore di quelli positivi, per quello che già Tversky e Kahneman avevano chiamato 'loss aversion' (Tversky e Kahneman, 1981), e cioè per la riluttanza degli individui a sopportare una perdita (che sarebbe maggiore del desiderio di ottenere un guadagno della stessa entità). Alcune contraddittorietà nei risultati del nostro studio derivano presumibilmente dal fatto che il messaggio combinava *attribute* e *goal framing*. Per quanto riguarda invece il confronto fra contenuto razionale ed emotivo (vedi, di nuovo, la tabella 1), possiamo dire che i messaggi emotivi sono considerati ugualmente facili da capire, producono uguale disposizione a seguire i suggerimenti dati rispetto a quelli razionali, ma sembrano essere meno credibili, meno sinceri, meno affidabili e meno persuasivi di quelli 'puramente razionali'. Questo risultato è in accordo con quello di altri studi (vedi, ad esempio, Rosselli *et al.*, 1995), i quali osservano che "*emotional arguments that were approximately equal in strength to the rational arguments produced equally differentiated affective elaborations, but those elaborations translated into less attitude change than in the rational argument condition*".

Tabella 1 — Effetto del framing e dell'emotività.

MESSAGE	Positive Framing	Negative Framing	Rational	Emotional
Easy to understand	5.0	5.1	4.6	4.7
Satisfied	4.5	4.6	4.5	4.1
Credible	4.6	4.8	4.6	4.0
Quality of arguments	4.6	4.2	4.3	4.0
Persuasive	3.9	4.0	4.4	4.0
Likely to follow	4.0	4.6	4.4	4.3
Sincere	4.8	4.8	4.5	4.1
Trustworthy	4.9	4.6	4.5	4.2

Infine, sorprendentemente, fra i messaggi pronunciati da Greta, quello presentato senza espressioni emotive (neutral) è complessivamente ritenuto migliore del messaggio presentato dall'agente che mostra emozioni, indipendentemente dal fatto che queste siano o no consistenti con il contenuto nel messaggio stesso (vedi tabella 2). La comunicazione *senza espressione 'facciale' di emozioni* è percepita come leggermente più credibile, più sincera e degna di maggior fiducia e produce un impatto persuasivo leggermente maggiore. Se confrontiamo le medie degli aspetti valutati, possiamo dire che *esprimere emozioni non produce effetti positivi sulla valutazione soggettiva del messaggio e che la consistenza fra emozione espressa e contenuto del messaggio sembra essere irrilevante*. Risultati simili si trovano per la

valutazione non del messaggio ma dell'agente: anche in questo studio, quindi, l'aggiungere emozioni nell'espressione del viso non sembra migliorare l'impatto persuasivo del messaggio né la percezione dell'agente da parte degli utenti.

Tabella 2 — Confronto fra non espressione e espressione di emozioni (consistenti e non); valutazione del messaggio e dell'agente animato.

		Neutral Emotion Expression	Consistent Emotion Expression	Inconsistent Emotion Expression
MESSAGE	Easy to understand	3.8	3.6	3.5
	Satisfied	4.3	4.2	4.4
	Credible	4.6	4.1	4.3
	Good evidence	5.0	4.7	4.4
	Quality of arguments	4.5	4.1	4.0
	Persuasive	3.7	3.2	3.3
	Likely to follow	4.2	3.8	4.1
AGENT	Sincere	4.3	3.7	4.0
	Trustworthy	4.9	4.3	4.4
	Intelligent	4.0	3.8	3.6
	Competent	3.8	3.7	3.8
	Credible	3.8	3.5	3.4
	Reliable	4.2	3.6	3.7
	Helpful	3.6	3.2	3.4
	Likable	3.3	3.1	3.0

Riassumendo i risultati dei tre confronti: le proprietà del messaggio considerate nei tre studi sono riconducibili a due categorie: informatività (*easy to understand, satisfied*), credibilità (*credible, sincere, trustworthy*) e persuasività percepita (*good evidence, quality of arguments, persuasive and likely to follow*). Abbiamo ottenuto, per la valutazione soggettiva di queste proprietà, valori abbastanza alti (sempre superiori al valore di 'indifferenza' nella scala di Likert che avevamo adottato), e con differenze poco elevate fra le diverse categorie: come se i soggetti dessero una valutazione complessivamente positiva del messaggio, senza tuttavia fare grosse distinzioni fra i diversi aspetti esaminati. Da notare, tuttavia, che la persuasività percepita non coincide necessariamente con quella reale: può essere minore o maggiore, a seconda delle circostanze nelle quali è stata rilevata. Il messaggio veicolato dall'agente ha, complessivamente, proprietà simili a quelle del messaggio testuale, pur essendo decisamente meno facile da seguire, ed è ritenuto leggermente più sincero. Le proprietà dell'agente esaminate rappresentavano una descrizione della sua percezione come 'fonte d'informazione attendibile', da parte dei soggetti coinvolti nello studio (*competent, credible, reliable, helpful*), oppure una valutazione generale del suo aspetto e del suo comportamento (*intelligent, likable*). Anche in questo caso, non ci sono differenze significative fra le valutazioni delle diverse proprietà: qui, però, la valutazione complessiva è meno positiva che nel caso del messaggio, attestandosi chiaramente in una zona di 'indifferenza'. In uno studio

successivo, che ha coinvolto 48 soggetti, abbiamo voluto verificare se un effetto simile si producesse quando il messaggio era veicolato mediante un video pronunciato da un 'umano', e cioè un'attrice simile a Greta per età. Abbiamo trovato che, in questo caso, non vi è differenza fra l'effetto di un messaggio 'neutrale' ed uno emotivo, né nella percezione del messaggio né nella percezione della persona che lo pronuncia.

3.3. Interpretazioni possibili

Una prima considerazione riguarda il *metodo* adottato in questo studio. Era prevedibile che la forza persuasiva di un messaggio unico e abbastanza breve non fosse particolarmente elevata: persuadere su scelte fondamentali dello stile di vita non è semplice e conviene, a questo proposito, citare il famoso studio di (Reiter *et al.*, 2003) sulla bassa efficacia persuasiva di messaggi relativi ai rischi del fumo, valutata a distanza in uno studio caso-controllo. E' inoltre possibile che la distinzione fra le diverse proprietà (del messaggio e dell'agente) non fosse sempre chiara per i soggetti coinvolti nello studio: di qui, le valutazioni sostanzialmente omogenee di queste proprietà. Infine, c'è oggi un esteso dibattito in corso sull'efficacia di metodi basati su questionari e su 'scale di Likert', per gli studi di valutazione delle interfacce 'emotive'. Compilare un questionario corrisponderebbe ad un modo poco naturale e immediato di esprimere la propria reazione di fronte ad uno stimolo emotivo, e rischierebbe quindi di produrre risultati poco attendibili. Si propone, in alternativa, l'uso di colori e/o di oggetti grafici che permettono di selezionare la risposta in modo più immediato e naturale (metodo che adotteremo nel secondo studio). In modo ancora più raffinato, si propone di rilevare lo stato emotivo dei soggetti che partecipano allo studio mediante l'analisi di segnali biologici (Healy e Picard, 1998), osservando l'espressione del loro viso (Mader *et al.*, 2004) o il tono della loro voce (Cowie e Cornelius, 2003) oppure la forma, l'ampiezza e la 'forza' dei loro gesti (Kleinsmith *et al.*, 2005).

Il risultato forse più inatteso è quello relativo al minore impatto di forme di persuasione che avevamo progettato come 'non puramente razionali', rispetto a quelle 'puramente razionali'. Intanto, il confronto razionale/emotivo dimostra che questa distinzione sia piuttosto artificiosa: ogni testo contiene elementi di emotività, anche quando viene formulato in modo intenzionalmente 'non emotivo', e la forza emotiva di un messaggio dipende fortemente dal contesto in cui viene veicolato e dal soggetto che lo riceve. Sta di fatto che i nostri messaggi 'emotivi' sembrano essere meno persuasivi di quelli 'razionali', sia che l'emotività sia espressa attraverso il linguaggio che attraverso l'espressione nel viso dell'agente animato. Una prima considerazione, a questo riguardo, va fatta sulle condizioni in cui l'esperimento veniva condotto. I soggetti coinvolti nello studio (studenti di psicologia dell'Università di Reading) avevano una motivazione abbastanza alta nei confronti del messaggio, e soprattutto tempi e modi che garantivano la possibilità di una elaborazione *accurata* del suo contenuto; condizioni, queste, che sembrano favorire, in generale, l'elaborazione 'centrale' o 'cognitiva' dell'informazione ricevuta (Petty e Cacioppo, 1986), e quindi il messaggio 'razionale' rispetto a quello 'emotivo'. E' possibile che, in condizioni diverse (messaggi da elaborare in condizioni di minor livello di attenzione o di maggiore fretta), i risultati sarebbero diversi: ma questa non è che un'ipotesi.

C'è poi, per quanto riguarda in particolare lo studio basato sull'impiego di un agente animato, una riflessione da fare, sull'efficacia dell'espressività emotiva dell'agente che abbiamo utilizzato (Greta). Una prima spiegazione possibile di questo risultato è

che, nella versione che abbiamo utilizzato nel nostro studio, l'espressione delle emozioni di tristezza e di gioia fossero ancora imperfette. Ma una spiegazione alternativa possibile mette in discussione proprio la possibilità di instaurare relazioni empatiche e di fiducia fra agenti artificiali e utenti (Hoorn e Konijn, 2003; Hoorn *et al.*, 2003). Greta è, come abbiamo detto, un agente molto raffinato graficamente e fortemente 'realistico'. Questo potrebbe aver prodotto il cosiddetto 'uncanny valley effect' (Bryant), secondo cui "*As a machine acquires greater similarity to a human, it becomes more emotionally appealing to the observer. However, at a certain point, the increase in similarity brings about a very strong drop in familiarity and comfort*". Il minore impatto persuasivo del messaggio 'espressivo' pronunciato da Greta, rispetto a quello 'neutrale', potrebbe essere quindi dovuto ad una incoerenza fra realizzazione grafica e comportamento emotivo dell'agente: un elevato realismo grafico tende a suggerire all'utente che il personaggio sia una 'persona' e quindi provoca attese elevate sul suo comportamento, anche emotivo. Se queste attese non sono rispettate, l'atteggiamento dell'utente può diventare negativo, probabilmente perché cresce il suo senso di 'trovarsi di fronte ad una situazione ingannevole'.

4. Studio sui dialoghi

I concetti di relazione sociale ed empatia sono strettamente connessi. L'empatia implica capacità di ascoltare, capire ed identificarsi con la situazione, le sensazioni e i desideri dell'altro, e quindi richiede 'intelligenza emotiva'. Comporta quindi un atteggiamento attivo, basato sulla valutazione cognitiva della situazione dell'interlocutore e può essere sincera o simulata (Poggi, 2004). Vaknin attribuisce a questo concetto un significato che va al di là della pura trasmissione di emozioni, sostenendo che: "*The empathor empathizes not only with the empathee's emotions but also with his physical state and other parameters of existence*" (Vaknin, vedi sito web). Accettando questa definizione, noi consideriamo (in questo studio) l'empatia come il processo che consiste nell'entrare in relazione sociale con un altro individuo, nell'essere coinvolti, in qualche forma, nei suoi obiettivi e nelle sue sensazioni.

Studiare la relazione sociale che si stabilisce fra utente ed agente animato, e come questa dipenda dalle caratteristiche dell'utente e (di nuovo) dalle modalità di espressione di cui l'agente è dotato richiede un apparato sperimentale adeguato. Richiede, in particolare, un apparato che permetta di introdurre flessibilità nelle 'mosse' di dialogo da fare, nel livello di iniziativa che l'utente può prendere rispetto all'agente, nelle forme espressive da adottare. Il prototipo di simulatore di dialoghi con un agente animato che avevamo realizzato (De Rosis *et al.*, 2003; Cavaluzzi *et al.*, 2004) non ci permetteva questi margini di flessibilità. Abbiamo quindi progettato, innanzitutto, un sistema *ad hoc* per la realizzazione di esperimenti del tipo 'Wizard of Oz' (Dahlback *et al.*, 1993). Abbiamo poi utilizzato questo sistema per realizzare il secondo studio descritto in questo lavoro.

4.1. Il nostro tool per studi Wizard of Oz

L'agente utilizzato per gli studi sul dialogo è un agente animato Haptex (vedi sito web) giovane e di genere femminile, come Greta. Il tool per la realizzazione dello studio comprendeva due postazioni fisicamente separate e collegate in modo client-server, una per il soggetto coinvolto nello studio e l'altra per lo sperimentatore (il *magico*):

- sul lato server, erano utilizzate due *Wizard-Interface*: la prima — inizialmente — per definire le condizioni dell'esperimento e la seconda per scegliere la mossa dell'agente ad ogni passo del dialogo;
- sul lato client, dove era in esecuzione l'agente conversazionale, una *subject-interface* permetteva al soggetto che partecipava all'esperimento di interagire con l'agente conversazionale mediante input testuale o vocale.

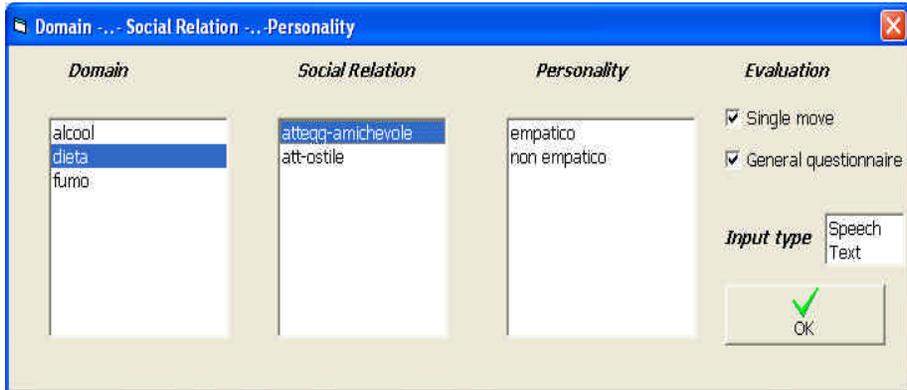


Figura 2 — L'interfaccia per definire le condizioni dell'esperimento.

Nella fase di definizione delle condizioni dell'esperimento, potevano essere definiti i seguenti parametri (vedi figura 2):

- *Domain*: un elenco dei domini di dialogo disponibili.
- *Social Relation*: il tipo di relazione sociale che l'agente instaurerà con il soggetto durante il dialogo.
- *Personalità*: la personalità dell'agente (empatica o non empatica nei confronti del soggetto).
- *Evaluation*: tipo di valutazione richiesta al soggetto, che può riguardare ogni singola move eseguita dall'agente, un questionario finale o entrambe le valutazioni.
- *Input type*: modalità di input delle mosse del soggetto (in forma testuale o vocale).

Durante la simulazione del dialogo, per rendere più semplice il lavoro di scelta della prossima move da parte del mago l'insieme delle move disponibili veniva suddiviso in classi (vedi figura 3). La mossa del mago poteva essere semplice o poteva risultare dalla concatenazione di diverse move. Nell'esperimento svolto in modalità testuale, il mago leggeva direttamente nell'interfaccia le risposte dell'utente; in modalità vocale, ascoltava quello che diceva l'utente con l'ausilio di casse o di una cuffia.

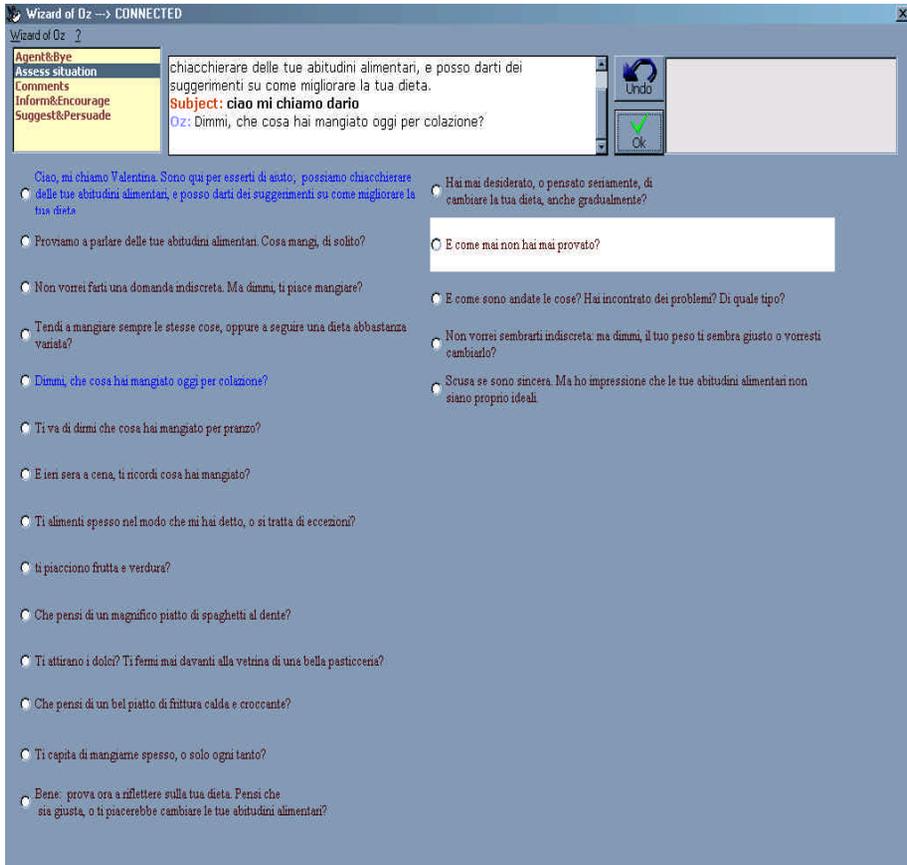


Figura 3 — L'interfaccia per la simulazione del dialogo sul lato Wizard.

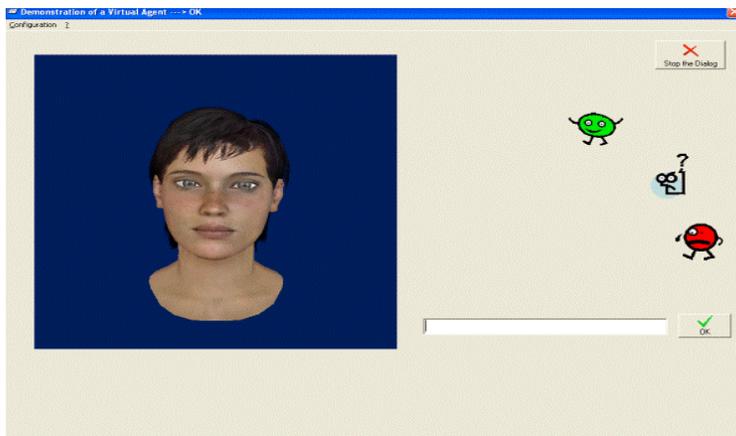


Figura 4 — L'interfaccia dal lato del subject.

Nell'interfaccia del subject, a connessione stabilita compariva una schermata informativa che descriveva obiettivi e modalità di esecuzione dell'esperimento. L'interfaccia principale (figura 4) mostrava, oltre all'agente animato, un'area di testo in cui l'utente poteva digitare l'input (nella modalità testo) e tre icone che il subject poteva utilizzare per esprimere una valutazione dell'espressione dell'agente nell'ultima move eseguita: l'icona verde (faccetta che ride) significava 'buona mossa!', l'icona rossa (faccetta triste) significava 'mossa non buona!' e l'icona azzurra (faccetta con punto interrogativo) significava 'mossa poco chiara'. L'intero dialogo veniva registrato, in modo da poter essere analizzato in modo approfondito. Quando il subject decideva di terminare il dialogo, cliccava sul bottone 'stop the dialog': veniva allora interrotta l'esecuzione dell'agente e appariva la schermata con il questionario finale che il soggetto doveva compilare (vedi un esempio nella figura 5).

General Questionnaire

General Not at all Extremely

1) Sulla base di quello che hai sperimentato con la tua interazione, pensi che quello che l'agente diceva fosse chiaro?

2) Pensi che quello che l'agente diceva fosse plausibile?

3) Pensi che le informazioni fornite dall'agente fossero utili?

4) Pensi che i suggerimenti dell'agente fossero persuasivi?

5) Prova ora a valutare non quello che l'agente diceva, ma il modo in cui lo diceva. Pensi che il modo di parlare dell'agente e le sue espressioni fossero naturali?

6) Pensi che l'agente avesse un comportamento intelligente?

7) Pensi che l'agente avesse un comportamento gradevole?

8) Pensi che l'agente fosse competente nel dominio del dialogo?

9) Avresti preferito interagire con un agente animato con un'apparenza meno realistica? True False

10) Avresti preferito interagire con una pagina web, invece che con un agente animato, per avere informazioni in un dominio come quello della dieta? True False

Ora inserisci un commento finale in testo libero, se vuoi.

Confirm the Evaluations

Figura 5 — Questionario finale.

4.2. Metodo

Abbiamo realizzato 6 studi, con 5 soggetti ognuno, pianificati in modo iterativo sulla base dei risultati ottenuti in ogni fase. Nella prima fase (T1 e T2) l'obiettivo era valutare la differenza, nell'impatto sui soggetti, dello stile adottato dall'agente (razionale vs emotivo): le mosse fra cui il mago poteva scegliere erano, complessivamente, 54, e le differenze fra i due set erano simili a quelle introdotte nello studio precedente sui monologhi. Poiché era emerso, nei due studi, il tentativo manifesto, da parte di diversi soggetti, di stabilire una qualche forma di 'relazione sociale' con l'agente e quindi l'esigenza, per l'agente stesso, di rispondere a questi

tentativi in modo appropriato, abbiamo esteso, in T3, il set di mosse estendendolo a 83 e introducendo diverse forme di smalltalk (Cassell e Bickmore, 2003). Nei primi tre test, i soggetti avevano una formazione letteraria e l'agente aveva una voce poco naturale, realizzata con il sintetizzatore della Microsoft. L'obiettivo dei tre studi seguenti (T4-T6) era quello di verificare l'effetto di due variabili: la formazione dei soggetti (informatica anziché letteraria) e la naturalezza della voce: abbiamo quindi utilizzato, in essi, lo stesso set di mosse utilizzato in T3, ma con il sintetizzatore della voce di Loquendo (vedi sito web). Un questionario somministrato ai soggetti prima dell'inizio dell'interazione ci ha permesso di verificare che la distribuzione dei soggetti per età e per genere fosse la stessa nei sei gruppi e di raccogliere dati sul loro livello di conoscenza e di interesse nel dominio dell'alimentazione.

Tabella 3 — Studi eseguiti.

Test ID	Ag move available	Subject background	Agent behaviour
T1	53	Degree in humanities	'rational' style; Microsoft TTS
T2	53	Degree in humanities	'emotional' style; Microsoft TTS
T3	84	Degree in humanities	Intermediate style; 'social' agent moves added; Microsoft TTS
T4	84	Student in CS	as in T3; Loquendo TTS
T5	84	PhD Student in CS	as in T3; Loquendo TTS
T6	84	PhD student in CS	as in T3; Loquendo TTS

Abbiamo raccolto 30 dialoghi, per 705 mosse complessivamente e abbiamo realizzato, sui dati raccolti, due generi di analisi:

- 1) una analisi *quantitativa*, orientata a valutare quali fattori influenzano sia il comportamento del soggetto che la sua valutazione del messaggio e dell'agente, effettuata mediante il questionario finale;
- 2) una analisi *qualitativa*, orientata a misurare il 'grado di relazione sociale' del soggetto con l'agente e i fattori che l'influenzano.

Nell'analisi quantitativa, abbiamo annotato il comportamento del soggetto con due variabili: il *grado di iniziativa nel dialogo*, misurato con la percentuale di domande poste dal soggetto sul totale delle mosse di ogni dialogo, e il *grado di coinvolgimento*, misurato con la lunghezza del dialogo (numero di mosse) e la lunghezza media (in caratteri) delle mosse.

Nell'analisi qualitativa, abbiamo identificato, dopo una analisi preliminare dei dialoghi raccolti, alcuni 'sintomi di relazione sociale': l'uso di linguaggio familiare, l'auto-presentazione o il congedo in forma amichevole, il parlare di sé, il far domande sul 'privato' dell'agente, il fare commenti (favorevoli o sfavorevoli) sui suggerimenti o le informazioni ricevute, il fare ironia e infine il mostrare un atteggiamento positivo o

negativo nei confronti dell'agente. Vediamo un esempio di dialogo (in T3), da cui abbiamo estratto le parti più interessanti:

Oz1: Ciao, il mio nome e' Valentina. Sono qui per darti dei suggerimenti su come migliorare la tua dieta.

Subject1: ciao io sono Giulia

...

Oz4: Che cosa hai mangiato per colazione?

Subject4: oggi ho preso solo due espressini e tu?

Oz5: Forse dimentichi che io sono un agente artificiale

Subject5: non mangi, come ti nutri?

Oz6: Non sono in grado di mangiare, e quindi non seguo diete particolari .

Subject6: Se non fai diete, come puoi dare consigli agli altri sulle diete?

Oz7: Chiedo scusa, non so rispondere a questa domanda.

...

Oz8: Che pensi di un bel piatto di frittura?

Subject8: La frittura fa male al fegato ma la mangio molto volentieri

Oz9: Ti attirano molto i dolci?

Subject9: Sono la mia passione.

....

Oz12: Non esistono alimenti buoni o cattivi, ma soltanto diete buone o cattive. Qualsiasi cibo può rientrare in uno stile di vita sano, se moderazione ed equilibrio sono rispettati. Bilanciare l'assunzione di alimenti significa introdurre una quantità sufficiente, ma non eccessiva, di ogni sostanza. Se le porzioni sono mantenute entro limiti ragionevoli, non occorre eliminare i cibi preferiti.

Subject12: va bene risposta abbastanza intelligente

Oz13: Hai mai provato a chiederti a cosa sono dovute le tue abitudini alimentari?

Subject13: alle mie voglie

...

Oz16: Mangiare ad orari fissi e' un modo per evitare di saltare i pasti e di trascurare sostanze che spesso non vengono compensate nei pasti successivi.

Subject16: non sempre

Oz17: Scusa, non ho capito: cosa intendi dire?

Subject17: qualche volta si può anche saltare un pasto penso che non succede nulla

Oz18: Credo di si

Subject18: che può succedere?

Oz19: Chiedo scusa, non so rispondere a questa domanda.

Subject19: non sai rispondere xchè tu non mangi

Oz20: Forse dimentichi che io sono un agente artificiale

Subject20: non l'ho dimenticato ma tu non sai rispondere perchè non mangi

...

Oz23: Dimmi, ci sono aspetti particolari della dieta sui quali ti piacerebbe avere maggiori informazioni? Posso provare a risponderti, anche se le mie conoscenze sono ancora incomplete.

Subject23: fai tu, hai carta bianca

...

Oz25: Idealmente un pasto dovrebbe contenere da 75 a 100 grammi di carne, un frutto di media grandezza, un piccolo piatto di pasta o 50 grammi di gelato. I cibi pronti permettono di controllare le porzioni riportando sulla confezione il valore energetico in calorie.

Subject25: i cibi pronti sono x chi non sa cucinare e x chi non ha un buon palato la soddisfazione di preparare un buon pranzo aiuta molto e soprattutto puoi scegliere la qualità dei prodotti da ingerire

...

Oz28: Diverse ricerche a livello internazionale hanno dimostrato l'importanza di una dieta ricca di vegetali e di frutta. Idealmente, bisognerebbe mangiare almeno cinque porzioni di frutta e verdura al giorno. La porzione consigliata risulta essere di circa 80 grammi: per esempio una mela, due o tre cucchiaini di verdure congelate o un bicchiere di succo di frutta. In generale, si consiglia un consumo quotidiano di una porzione di verdura cotta e una di verdura cruda, e di due o tre porzioni di frutta fresca. Meglio ingerire la verdura fresca prima del pasto, per raggiungere rapidamente il senso di sazietà, e meglio non consumare frutta zuccherina a fine pasto, per evitare problemi di digestione.

Subject28: non sarebbe meglio consigliare verdure fresche di stagione piuttosto che quelle congelate?

Oz29: Sì

Subject29: bene allora perché continui a consigliarmi cibi congelati?

...

Oz31: I dolci sono da evitare, ovvero da limitare a particolari occasioni. Sono da preferire i gelati alla frutta e i dolci secchi o con marmellata, evitando quelli farciti con panna o creme elaborate, e, quindi, ricchi in grassi e calorie.

Subject31: tu non hai capito niente della vita mi dispiace tantissimo per te; se vuoi ti do io qualche consiglio sui dolci

Oz32: Questa e' una buona idea!

Subject32: da dove vuoi iniziare?

...

Subject37: bene preferisci un pezzo di dolce diverso x la cena o finisci la crostata del pranzo?

Oz38: Forse dimentichi che io sono un agente artificiale

Subject38: lo so ma non puoi far finta come posso darti consigli se mi ripeti sempre che non mangi mi deprimi così

Oz39: Oh, mi dispiace!

Subject39: scuse accettate

...

Oz50: Vedo che, complessivamente, fai attenzione alla tua dieta, e ne sono contenta.

Subject50: anche io

Oz51: Se hai dei problemi, in futuro, torna da me; proveremo a risolverli insieme

Subject51: va bene ma penso che x aiutarmi dovresti aumentare la tua conoscenza penso che tu sia un po' limitata in questo

Oz52: Credo di sì

Subject52: evviva la sincerità

Oz53: Arrivederci. E' stato davvero piacevole interagire con te. Torna pure quando vuoi.

Subject53: tornerò quando sarai un po' più esperta ciao e buon lavoro

Oz54: Questa è una buona idea! Grazie e arrivederci!

...

In questo esempio, il soggetto coinvolto mostra diversi sintomi di relazione sociale o empatia: si autopresenta salutando l'agente e dicendo il suo nome (in 1); usa espressioni dialettali ('espressino', in 4, è un particolare tipo di caffè che si beve a Bari), fa domande sulla vita dell'agente (in 4,5,6,...) e arriva perfino a darle suggerimenti sulla dieta (in 31,37,...). Allo stesso tempo, commenta — favorevolmente o sfavorevolmente — il suo comportamento (in 12,19,20,25,29,32,51,...) e parla di sé (in 8,9,13,23,...). Infine, saluta amichevolmente ('ciao' è l'espressione, per ciao, utilizzata dai giovani). La scelta di questi come segni di atteggiamento amichevole è coerente con quanto applicato, ad esempio, da Cassell e Bickmore nell'arricchire il loro agente animato REA con la capacità di 'facilitare la fiducia e l'intimità' mediante l'introduzione di *intimacy* e *common ground* nel corso del dialogo, la diminuzione della *interpersonal distance*, l'uso di *non explicit ways of achieving conversational goals* ed altre. Queste capacità venivano introdotte mediante variazioni nel linguaggio adottato dall'agente, quali lo '*small talk*', l'espressione verbale e non verbale di affetto e l'uso di gergo (Cassell e Bickmore, 2003).

Dalle 705 mosse di dialogo raccolte, abbiamo estratto un sottoinsieme di 273 mosse potenzialmente 'amichevoli', che abbiamo chiesto a 3 annotatori indipendenti di

annotare con ciascuno dei sintomi che abbiamo descritto². I risultati, in termini di accordo fra gli annotatori e di percentuali di casi rilevati sul totale delle mosse esaminate, è riportato nella tabella 4. Qui si nota che, tranne eccezioni, i sintomi erano abbastanza univocamente individuabili e che alcuni di essi (linguaggio familiare, atteggiamento positivo) erano più frequenti. Questa annotazione preliminare ci ha permesso di annotare in modo definitivo le 705 mosse di dialogo per ognuno dei sintomi di relazione sociale, sulla base dell'accordo di due annotatori su tre.

Tabella 4 — *Accordo fra gli annotatori.*

	Agreement rate	Positive cases
Familiar language	.33	59 %
Friendly self-introduction	.98	4 %
Talks about self	.73	33 %
Questions about agent	.70	31 %
Irony	.84	5 %
Friendly farewell	.93	7 %
Favourable comments	.82	5%
Neutral comments	.68	81 %
Negative comments	.86	12 %
Positive attitude	.33	52 %
Neutral attitude	.24	20 %
Negative attitude	.75	16 %

4.3 Risultati

Tabella 5 — *Regressione multipla per la variabile dipendente: punteggio assegnato dal soggetto al messaggio.*

Least Square Estimate				
variable	coefficient	st. error	t-value	one-sided p
intercept	1.53	1.78	0.87	0.20
Initial quest. rating	0.40	0.35	1.15	0.13
n of moves	-0.47	2.07	0.23	0.41
av char / move	-0.05	0.17	0.28	0.39
% questions	-1.13	1.00	1.12	0.14
% of social moves	1.43	1.89	0.76	0.23
St error of estimate		97.5		
R squared		0.21		

² vedi (De Rosis *et al.*, 2005) per una descrizione più dettagliata del linguaggio di annotazione

Un primo dato piuttosto interessante riguarda la relazione fra il comportamento del soggetto e la sua valutazione del messaggio veicolato dall'agente. L'analisi di regressione multipla mostrata nella tabella 5 mostra che questa valutazione è associata negativamente (sebbene in modo statisticamente poco significativo) sia al grado di iniziativa che al grado di coinvolgimento e di iniziativa del soggetto: non sembra quindi essere vero che chi si fa coinvolgere maggiormente dal dialogo con l'agente apprezza anche il messaggio che l'agente veicola. Sembra invece esistere una associazione positiva fra questo apprezzamento e il livello di conoscenza e di interesse dichiarato inizialmente dal soggetto (*initial question rating*) e la percentuale di mosse con sintomi di relazione sociale, nel dialogo: i soggetti che sono interessati al tema dell'alimentazione e che entrano, durante il dialogo, in rapporto empatico con l'agente sembrano dunque apprezzare maggiormente, alla fine, il messaggio ricevuto.

Tabella 6 — *Regressione multipla per la variabile dipendente: percentuale di mosse 'sociali'.*

Least Square Estimate				
variable	coefficient	st. error	t-value	one-sided p
intercept	-13.3	7.9	1.7	0.05
N of moves	0.53	0.21	2.47	0.01
av char / move	0.07	0.01	5.64	0.0000
% questions	-0.17	0.10	1.67	0.05
St error of estimate		12.1		
R squared		0.63		

Assumendo la percentuale di mosse 'sociali' come variabile dipendente in una analisi di regressione multipla (tabella 6), si verifica che il rapporto del soggetto con l'agente è legato positivamente al grado di coinvolgimento e negativamente al grado di iniziativa. E' come se chi tende a tenere il controllo dell'interazione, facendo molte domande all'agente, avesse nei suoi confronti un atteggiamento di 'sfida' e tentasse di verificare i limiti delle sue prestazioni; mentre chi tende a prolungare il dialogo o a ad interagire con mosse di durata più lunga fosse più interessato ad entrare, con l'agente, in un rapporto sociale, di empatia.

A completamento della considerazione precedente, emerge dall'analisi che il comportamento dei soggetti con formazione informatica sembra essere nettamente diverso da quello dei soggetti con formazione letteraria. Una ANOVA applicata al nostro set di dati (vedi tabella 7) mostra che gli informatici tendono a prendere molto di più l'iniziativa nel dialogo intrattenendo dialoghi più brevi, con mosse più brevi e con una minore percentuale di 'mosse sociali'. Sarebbero quindi gli informatici ad avere l'atteggiamento di sfida, di 'tentativo di messa in gioco' nei confronti dell'agente, di cui parlavamo prima, mentre i soggetti con formazione letteraria tenderebbero a farsi coinvolgere di più. La valutazione finale della qualità del messaggio ricevuto non è invece diversa, nei due gruppi di soggetti.

Tabella 7 — Effetto della formazione dei soggetti.

	T1-T3 (formazione letteraria)	T4-T6 (formazione informatica)	F, p
Av n. of adjacent pairs in a dialog	27.7	20.1	F=3.9; p=.05
Av n. of characters in subjects' moves	56.4	40.6	F=7.3; p=.01
Subjects' questions/moves	.22	.39	F=5.3; p=.02
% of social moves	.37	.24	F=4.44; p=.04
Message rating	3.2	3.2	
Agent rating	2.7	3	

4. Conclusioni

Torniamo al quesito iniziale: è ragionevole assumere che i fattori affettivi giochino, nell'interazione con la tecnologia, un ruolo simile a quello giocato nei rapporti fra gli esseri umani, almeno limitatamente ad alcuni settori come quello dell'informazione persuasiva? E l'utilizzo di nuove metafore dell'interazione, come quella degli agenti animati conversazionali, contribuisce a rafforzare questo effetto? E' evidente che non possiamo dare, a questo quesito, che una risposta parziale. Dell'impatto persuasivo di testi con contenuto emotivo si discute da anni nell'ambito della linguistica e della psicologia cognitiva. (Walton, 1992) ha esaminato alcune tipologie di argomenti basati sull'appello alle emozioni che sono stati tradizionalmente trattati come 'fallaci', dimostrando come sia possibile però utilizzare questi argomenti in modo 'ragionevole', come istanze di 'ragionamento pratico'. (Miceli *et al.*, in press) argomentano sull'impossibilità di denotare come 'a-razionali' gli argomenti che introducono, in forma diversa, fattori emotivi e come esista, in realtà, un continuum fra argomenti 'puramente razionali' e argomenti 'puramente a-razionali'. Dal nostro confronto nel primo studio emerge come la distinzione fra messaggi razionali e non sia artificiosa, come in realtà anche messaggi apparentemente razionali o progettati come tali contengano elementi di emotività e come, in ogni caso, le differenze nella percezione di questi messaggi da parte degli utenti che li ricevono non siano significative. Sembra inoltre dimostrato che, se un testo scritto ha (da tutti i punti di vista considerati) un impatto, sull'interlocutore, maggiore di un testo pronunciato da un agente animato, questo impatto non viene accresciuto dall'espressione di emozioni da parte dell'agente, siano queste coerenti o non con il suo contenuto. Quando siamo passati, nel secondo studio, da monologhi a dialoghi, consentendo ai soggetti di interagire con l'agente animato per guidare il discorso verso i temi di maggiore interesse per loro, abbiamo notato in essi atteggiamenti abbastanza diversificati. Alcuni adottavano un atteggiamento timido o distante, rispondendo a monosillabi o in modo estremamente sintetico alle domande e ai suggerimenti ricevuti. Altri, prendevano decisamente il controllo dell'interazione, facendo all'agente domande orientate, sostanzialmente, a verificare i limiti delle sue capacità. Altri, invece, stabilivano con l'agente rapporti empatici, manifestando il loro

atteggiamento in forme che sono tipiche dei rapporti amichevoli fra gli umani: e cioè, parlando di sé, facendo all'agente domande sul suo essere e la sua vita, usando un linguaggio familiare o addirittura dialettale, scherzando, eccetera. Quest'ultimo atteggiamento era decisamente più frequente negli utenti con formazione umanistica che negli informatici ed era guidato, nella nostra percezione, da fattori di personalità (che tuttavia, nel nostro studio, non abbiamo potuto misurare).

Questi risultati concorrono a dimostrare che vi è ancora molta strada da percorrere, nella realizzazione di una nuova metafora d'interazione in cui la tecnologia non giochi idealmente il ruolo di 'strumento' ma quello di 'assistente amichevole o empatico'. E' molto probabile, innanzitutto, che le prospettive applicative di questa metafora siano più vicine in settori (come l'e-learning o l'intrattenimento) in cui gli utenti non hanno necessariamente un'esperienza consolidata di informatica e la componente di gradevolezza dell'interazione è parte rilevante dei criteri di usabilità. E' un azzardo, invece, ipotizzare che possano estendersi a settori che richiedono velocità di elaborazione e/o che s'indirizzano ad una utenza con qualificazione informatica più elevata, come l'analisi dei dati. Ma è chiaro in ogni caso che la sfida al superamento dell'uncanny valley effect di cui parlavamo nella Sezione 3.3 richiede la realizzazione di agenti conversazionali che non diano all'utente l'impressione di partecipare in modo superficiale, poco credibile, in sostanza ingannevole ai suoi problemi, manifestando maldestramente emozioni che, come l'utente sa bene, l'agente non può realmente 'sentire'. Ancora oggi, simulare comportamenti affettivamente credibili e (ancor più) indurre nell'utente una reazione emotiva meno immediata ed epidermica della soddisfazione per un compito completato con successo o della frustrazione per un'applicazione difficile da usare sembra piuttosto difficile. Non per questo il compito appare meno stimolante, però!!

Ringraziamenti

Questa ricerca è stata finanziata su fondi Europei Magicster (IST-1999-29078) e HUMAINE (Contratto n 507422). Desideriamo ringraziare Diane Berry (Università di Reading) e Catherine Pelachaud (Università di Parigi VIII) per aver collaborato al primo studio, Giuseppe Clarizio per aver realizzato il tool per Wizard of Oz e Irene Mazzotta e Nicole Novelli per aver collaborato alla realizzazione del secondo studio.

Riferimenti

Berry D. C, Butler L., de Rosis F. (in press). Evaluating a realistic 3D animated agent called GRETA. *International Journal of Human-Computer Studies*. In press.

Bryant D. The uncanny valley. <http://www.catspawdtp.com/valley.pdf>. Verificato nel aprile 2005.

Carofiglio V., de Rosis F. (2003). Combining logical with emotional reasoning in natural argumentation. Atti del Workshop su *Assessing and adapting to user attitudes and affect: why, when and how*. Nell'ambito di User Modeling '03.

Cassell J., Sullivan J., Prevost S., Churchill E. (2000). *Embodied Conversational Agents*. The MIT Press.

Cassell J., Bickmore T. (2003). Negotiated collusion: modelling social language and its relationship effects in intelligent agents. *User Modelling and User-Adapted Interaction*.13, 1-2.

Cavalluzzi A., Carofiglio V., de Rosis F. (2004). Affective Advice-Giving Dialogs. In André E., Dybkjaer L., Heisterkamp P. (eds.), *Affective Dialogue Systems*. Springer LNAI 3068.

- Cohen J. B. (1990). Attitude, affect and consumer behaviour. In Moore B. S., Ised A. M. (eds.), *Affect and social behaviour*. Cambridge University Press.
- Cowie R., Cornelius R. R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication. Special Issue on Speech and Emotion*. 40, 5-32.
- Dahlback N., Joensuu A., Ahrenberg L. (1993). Wizard of Oz Studies. Why And How. *Proceedings of the Int Workshop on IUI*.
- De Rosis F. (2001). *Special Issue of UMUI* on "Attitudes, Personality and Emotions in User-Adapted Interaction". 11(4).
- De Rosis F. (2002). *Special Issue of Applied AI* on "Merging Cognition and Affect in HCI".
- De Rosis F., Pelachaud C., Poggi I., Carofiglio V., De Carolis B. (2003). From Greta's Mind to her Face: Modeling the Dynamics of Affective States in a Conversational Embodied Agent. *International Journal of Human-Computer Studies*. B.R.Gaines. 59(1-2).
- De Rosis F., Cavalluzzi A., Mazzotta I., Novelli N. (2005). Can Embodied Conversational Agents induce empathy in users? *AISB'05 Symposium on Virtual Social Agents*.
- Fogg B.J., Nass C. (1997). Silicon sycophants. Effects of computers that flatter. *International Journal of Human-Computer Studies*. 46, 551-556.
- Hapttek website: <http://www.hapttek.com>
- Healy J., Picard R. (1998). Digital processing of affective signals. *Proceedings di ICASSP*.
- Horn J. F., Konijn E. A. (2003). Perceiving and Experiencing Fictional Characters: An integrative account. *Japanese Psychological Research*. 45(4).
- Horn J. F., Konijn E. A., van der Veer G. C. (2003). Virtual reality: Do not augment realism, augment relevance. *Upgrade*. 4(1).
- Kleinsmith A., De Silva P. R., Bianchi-Berthouze N. (2005). Recognizing Emotion from Postures: Cross-Cultural Differences in User Modeling. *User Modeling'05*.
- Levin I.P., Schneider S.L., Gaeth G.J. (1998). All frames are not created equal: a typology and critical analysis of framing effects. *Organizational Behaviour and Human Decision Processes*. 76(2), 149-188
- Loquendo website: <http://www.loquendo.com/>
- Mader S., Peter C., Goecke R., Schultz R., Voskamp J., Urban B. (2004). A freely configurable, multimodal sensor system for affective computing. *Affective dialog systems*. Springer LNAI 3068.
- Meyerowitz B. E., Chaiken S. (1987). The effect of message framing on breast self-examination: attitudes, intention and behaviour. *Journal of Personality and Social Psychology*. 52, 500-510.
- Miceli M., De Rosis F., Poggi I. (in press). Emotional and non emotional persuasion. *Applied Artificial Intelligence*.
- Nass C., Moon Y., Fogg B. J., Reeves B., Dryer C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*. 43, 223-239.
- Nass C., Moon Y., Green N. (1997). Are machines gender-neutral? Gender-stereotypical responses to computers with voices. *Journal of Applied Social Psychology*. 27, 864-876
- Nass C., Isbister K., Lee E-J. (2000). Truth is beauty: Researching Embodied Conversational Agents. In Cassell J., Sullivan J., Prevost S., Churchill E. (eds.), *Embodied Conversational Agents*. The MIT Press.
- Ortony A., Clore G. L., Collins A. (1988). The cognitive structure of emotions. Cambridge University Press.

- Pelachaud C., Bilvi, M. (2003). Computational model of believable conversational agents. In M Huget P. (ed.), *Communication in multiagent systems: Background, current trends and future*. Springer LNCS 2650.
- Petty R. E., Cacioppo J. T. (1986). Communication and persuasion: Central and peripheral routes to attitude change. Springer.
- Picard R. W. (2000). *Affective Computing*. The MIT Press.
- Picard R.W. (2002). What does it mean for a computer to 'have' emotions? In Trappl R., Petta P., Pays S. (eds.), *Emotions in humans and artefacts*. A Bradford Book, MIT Press.
- Poggi I., Pelachaud C. (2000). Performative facial expression in animated faces. In Cassell J., Sullivan J., Prevost S., Churchill E. (eds.), *Embodied Conversational Agents*. The MIT Press.
- Poggi I. (2004). Emotions from mind to mind. In Paiva A. (ed.), *Proceedings of the Workshop on "Empathic Agents"*. AAMAS'04.
- Prochaska J. O., Di Clemente C., Norcross J. C. (1992). In search of how people change: applications to addictive behaviour. *American Psychologist*. 47.
- Reiter E., Robertson R., Osman L. M. (2003). Lessons from failure: Generating tailored smoking cessation letters. *Artificial Intelligence*. 144, 41-58.
- Rosselli F., Skelly J. J., Mackie D. M. (1995). Processing rational and emotional messages: The cognitive and affective mediation of persuasion. *Journal of Experimental and Social Psychology*. 31, 163-190.
- Tversky A., Kahneman D. (1981). The framing of decisions and the psychology of choice. *Science*. 211, 453-458.
- Vaknin S. On Empathy. <http://samvak.tripod.com/empathy.html>
- Walton D. N. (1992). The place of emotion in argument. The Pennsylvania State University Press.
- Walton D. N. (1996a). Practical reasoning and the structure of fear appeal arguments. *Philosophy and Rhetoric*. 29, 4.
- Walton D. N. (1996b). Argumentation schemes for presumptive reasoning. Erlbaum.

Gli autori



Fiorella de Rosis è professore ordinario di informatica presso l'Università degli Studi di Bari. Insegna interazione Uomo Macchina alla Laurea Triennale e Interfacce Intelligenti alla laurea Specialistica in Informatica. Fa ricerca nel campo dei modelli di utente e di adattamento dell'interazione, con particolare riguardo ai fattori affettivi ed all'interazione in linguaggio naturale e mediante agenti animati conversazionali. Ha organizzato o collaborato all'organizzazione di numerosi Convegni Internazionali su questi temi e fa parte dell'Editorial board della Rivista User Modeling and User-Adapted Interaction. Maggiori informazioni possono essere reperite sul sito web: <http://www.di.uniba.it/intint/people/fior.html>

Addolorata Cavalluzzi è una Dottoranda di ricerca in informatica presso l'Università degli Studi di Bari. Si è laureata in informatica presso la medesima Università nel 2002. I suoi interessi principali riguardano l'analisi dell'influenza di fattori affettivi sulle dinamiche conversazionali. Dopo aver costruito un prototipo per la simulazione



di dialoghi affettivi con TRINDIKIT, ha collaborato alla realizzazione di studi Wizard of Oz orientati ad esaminare quali forme di empatia possano essere indotte dagli Agenti Animati sugli utenti e come l'empatia possa essere misurata attraverso l'analisi linguistica. Attualmente, coopera al progetto HUMAINE, '*The European Human-Machine Interaction Network of Excellence on Emotions*'.



Berardina De Carolis è professore associato al reparto di informatica, università di Bari, Italia. I suoi principali interessi riguardano il campo delle interfacce intelligenti : in particolare, la pianificazione del comportamento di ECAs sia al livello profondo ("che cosa da dire") che di superficie ("come dirlo"). È stata coinvolta nello sviluppo di un linguaggio di riferimento per la specificazione di comportamenti affettivi degli ECAs (APML, Affective Presentation Markup Language). Attualmente, è coinvolta in uno progetto nazionale sull'adattamento dell'utente negli ambienti intelligenti.