A GOOD GESTURE: EXPLORING NONVERBAL COMMUNICATION FOR ROBUST SLDSs

Beatriz López Mencía[1], Álvaro Hernández Trapote[1], David Díaz Pardo de Vera[1], Doroteo Torre Toledano[2], Luis A. Hernández Gómez[1], Eduardo López Gonzalo[1]

[1] Grupo de Aplicaciones del Procesado de Señales, E.T.S.I.T, Universidad Politécnica de Madrid

[2] Escuela Politécnica Superior, Universidad Autónoma de Madrid

ABSTRACT

In this paper we propose a research framework to explore the possibilities that state-of-the-art embodied conversational agents (ECAs) technology can offer to overcome typical robustness problems in spoken language dialogue systems (SLDSs), such as error detection and recovery, changes of turn and clarification requests, that occur in many human-machine dialogue situations in real applications. Our goal is to study the effects of nonverbal communication throughout the dialogue, and find out to what extent ECAs can help overcome user frustration in critical situations. In particular, we have created a gestural repertoire that we will test and continue to refine and expand, to fit as closely as possible the users' expectations and intuitions, and to favour a more efficient and pleasant dialogue flow for the users. We also describe the test environment we have designed, simulating a realistic mobile application, as well as the evaluation methodology for the assessment, in forthcoming tests, of the potential benefits of adding nonverbal communication in complex dialogue situations.

1. INTRODUCTION

Spoken language dialogue systems and embodied conversational agents are being introduced in a rapidly increasing number of Human-Computer Interaction (HCI) applications. The technologies involved in SLDSs (speech recognition, dialogue design, etc.) are mature enough to allow the creation of trustworthy applications. However, robustness problems still arise in concrete limited dialogue systems because there are many error sources that may cause the system to perform poorly [1].

At the same time, embodied conversational agents (ECAs) are gaining prominence in HCI systems, since they make for more user-friendly applications while increasing communication effectiveness. There are many studies on the effects –from psychological to efficiency in goal achievement– ECAs have on users of a variety of applications (see [2] and [3]), but still very few (see [4]) on the impact of ECAs in directed dialogue situations where robustness is a problem.

We propose looking into the effects of adding an ECA to a concrete spoken dialogue system, and the potential benefits this may have, particularly regarding various difficult dialogue situations already identified by various leading authors in the field ([5] and [6]). This paper outlines the main elements of a research framework we have designed for these purposes.

Our research group is paying particular attention to videotelephony applications. We may now consider incorporating ECAs onto the new visual channel. Videotelephony has its own peculiarities which it may be relevant to take into account when developing ECAs for them (for instance, screen space is more limited).

2. HOW ECAS CAN BE USEFUL

There are many nonverbal elements of communication in everyday life that are important because they convey a considerable amount of information and qualify the spoken message, sometimes even to the extent that what is meant is actually the opposite of what is said. Showing objects, types of behaviour, mood, reactions, emotions and pointing towards something in the referential context (deictic gestures) are some of the functions of nonverbal language, which carries a great amount of semantic content related to people's attitudes and intentions in interaction processes (see [7]).

ECAs offer the possibility to combine several communication modes such as speech and gestures, making it possible, in theory, to create interfaces with which human-machine interaction is much more natural and comfortable. Despite the fact that we are still a long way from understanding how best to incorporate nonverbal communication to improve human-machine dialogue, ECAs are already being employed to improve interaction (see [8]). These are some situations in which an ECA could have a positive effect:

- Efficient turn management: the body language and expressiveness of agents are important not only to reinforce the spoken message, but also, as Cassell points out [2], to regulate the flow of the dialogue.
- Improving error recovery: the process of recognition error recovery usually leads to a certain degree of user frustration (see [9]). ECA's may help reduce frustration, and by doing so make error recovery more effective [10]. Indeed, it is common, once an error occurs, to enter in an error spiral, in which the system is trying to recover, the user gets ever more frustrated, and this frustration interferes in the recognition process (since, for example, users often repeat their previous utterance in a way that the system is less likely to understand), making the situation worse [11].
- Correct understanding of the state of the dialogue: a common problem in dialogue systems is that the user doesn't know whether or not the process is working normally [12].

This sometimes leads the dialogue to error states that could be avoided. The expressive capacity of ECAs could be used to reflect the state the system takes the dialogue to be in.

Our research framework, comprising a dialogue and ECA behaviour scheme, has been designed with these typical spoken dialogue system problems in mind, so that we may study the effect of an ECA in a variety of dialogue situations. In the following section we identify the main stages the dialogue can go through and we associate ECA behaviour and presentation strategies with each of them so that we may have a starting point with which to test real user-dialogue system interactions.

3. DIALOGUE STAGES AND ECA BEHAVIOUR

A variety of studies have been carried out on behavioural strategies for embodied conversational agents ([13], [14], [15], [16], [17]), which deal with behaviour in hypothetical situations and in terms of the informational goals of each particular interaction (be it human-human or human-machine). On our part, we will focus our attention on the overall dynamics in dialogue systems, focussing specifically on typical robustness problems, and how to favour smooth sailing through the different stages of the dialogue. Certain situations, combined with user expectations, can be sources of frustration, which negatively affects system robustness, and need to be considered with particular care, from dialogue initiation (welcoming messages) to termination.

We draw from existing research undertaken to try to understand the effects different gestures displayed by ECAs have on people, and we apply this knowledge to a real dialogue system. In Table 1 we show the basic set of gestures we are using as a starting point. They are based mainly on gestures described in [2] and [14], and on recommendations in [5], [15], [16], [17] and [18], to which we have added some suggestions of our own.

3.1. Initiation

The first contact users have with the system is through a welcome message at dialogue initiation. The inclusion of an ECA at this stage "humanises" the system [19]. This may predispose the user to be more forgiving when a problem arises in the dialogue. Interaction with a humanlike character may, however, induce in the user false expectations about the system, since she will tend anthropomorphisise it and believe it can deal with social interactions on a level similar to that of humans [20].

This is a problem, first because once a user has such high expectations the system can only end up disappointing her, and secondly because the user will tend to use more natural (and thus complex) communication, which the system is unable to handle, and the experience will ultimately be frustrating.

On the other hand, especially in the case of new users, contact with a dialoguing animated character may have the effect that the user's level of attention to the actual information that is being given is reduced (see [21] and [22]). Thus the goal is to present a human-like interface that is, at the same time, less striking and thus less distracting at first contact, and one that clearly "sets the rules" of the interaction and makes sure that the user keeps it framed within the capability of the system.

We have designed a welcome gesture for our ECA based on the recommendations in [17], to test whether or not it fosters a sense of ease in the user and helps her concentrate on the task at hand. Playing with the zoom, the size and the position of the ECA on the screen may also prove to be useful to frame the communication better. For instance, the character might appear first at mid-distance, visible from the waist up, giving a general greeting, then we might move on to a close-up of the face to give a welcome message more specific to the system, and then zoom out to a full-body view where the character explains the tasks and options ahead for the user; we have added this gestural sequence to a hand wave and smile [17] for our initial tests (see Table 1).

3.2. Termination

As we mentioned in Section 2, ECAs may be employed to help improve the user's understanding of the state of the dialogue. At dialogue termination the user should understand that the dialogue is being closed. We have implemented a gesture for our ECA, taking the main part (nod and hand wave) from [14], and adding turning around and walking away movements, dimming the lights at the end (see Table 1).

3.3. Turn Management

Turn management involves two basic actions: taking turn and giving turn. Again, an ECA's gestures can be designed to help the user see that the system is either taking the turn or handing it to the user. And again, in Table 1 we show the corresponding ECA gestures we will start testing with.

Note that apart from the ECA gestures, we also play with zoom and light intensity: when it's the ECA's turn to speak the camera zooms-in slightly and the light becomes brighter, and when it's the user's turn the camera zooms out and the lights dim. The idea is that, hopefully, the user will associate each level of light intensity with each of the turn modes, and so know when she is expected to speak.

Poor turn management is a major source of errors in SLDSs. An ECA can display body language and expression to indicate turn requests and changes more clearly, reducing the amount of errors [2]. The following are some typical examples of problem situations together with further considerations about ECA behaviour that could help avoid or recover from them:

- The user is not aware that she can interrupt, or be interrupted by, the ECA (barge-in). This leads to a less efficient communication and may be testing for the user's patience. An obvious solution is to have the ECA openly remind the user that she may interrupt. Also, the ECA could take a receptive position while talking, to encourage the user to interrupt whenever she wishes. For this purpose we will try taking a fixed mid-distance position, looking directly at the user.
- The user tries to interrupt at a point at which the barge-in feature is not active (deactivating barge-in can be useful in situations in which it is crucial for the user to listen to the system). If this happens the system does not process what the user has said, and when the system finally returns to listening mode there is silence from both parts: the system expects input from the user, and the user expects an answer. Often both finally break the silence at the same time and the cycle begins again, or, if the system caught part of the user's utterance, a recognition error will most likely occur and the system will fall into a recognition error recovery subdialogue that the user does not expect. To help avoid such faulty events the ECAs

demeanour should indicate as clearly as possible that the user is not being listened to at that particular moment. Speaking while looking away, perhaps at some object, and absence of attention cues (such as nodding) are possible ways to show that the user is not expected to interrupt the ECA. Since our present dialogue system produces fairly short utterances for the ECA, we are somewhat limited as to the active strategies to build into the ECA's behaviour. However, there are at least three cues the user could read to realise that the system didn't listen to what she said. The first is the fact that the system carries on speaking, ignoring the user's utterance. Second, at the end of the system's turn the ECA will perform a specific give-turn gesture. And third, after giving the turn the ECA will remain still and silent for a few seconds before performing a waiting gesture (leaning back slightly with her arms crossed, shifting the body weight from one leg to another; see Table 1). In addition, if the user still remains silent after yet another brief waiting period the system will offer help. It will be interesting to see at which point users realise that the system didn't register their utterance.

• A similar situation occurs if the Voice Activity Detector (VAD) fails and the system doesn't capture the user's entire utterance, or when the user simply doesn't say anything when she is expected to ("no input"). Again, both system and user end up waiting for each other to say something. And again, the strategy we use is to have the ECA display a waiting posture.

• It can also happen that the user doesn't speak but the VAD "thinks" she did, perhaps after detecting some background noise (a "phantom input"). The dialogue system's reaction to something the user didn't say can cause surprise and confusion in the user. Here the visible reactions of an ECA might help the user understand what has happened and allow her to steer the dialogue back on track.

3.4. Recognition Confidence Scheme

Once the user utterance has been recognised, information confirmation strategies are commonly used in dialogue systems. Different strategies are taken depending on the level of confidence in the correctness of the user locution as captured by the speech recognition unit (see [18]). Our scheme is as follows:

• High confidence: if recognition confidence is high enough to safely assume that no error has occurred, the dialogue strategy is made more fluent, with no confirmations being sought by the system.

• Intermediate confidence: the result is regarded as uncertain and the system tries implicit confirmation (by including the uncertain piece of information in a question about something else.) This, combined with a mixed initiative approach, allows the user to correct the system if an error did occur.

• Low confidence: in this case recognition has probably failed. When this happens the dialogue switches to a more guided strategy, with explicit confirmation of the collected information and no mixed initiative –*i.e.*, the user has no freedom and can only follow the instructions given by the system. The user's reply may confirm that the system understood correctly, in which case the dialogue continues to flow normally, or, on the other hand, it may show that there was a recognition error. In this case an error recovery mechanism begins.

In addition to the dialogue strategies, ECAs could also be used to reflect in their manner the level of confidence that the system has understood the user, in accordance with the confirmation dialogue strategies

While the user speaks, our ECA will, if the recognition confidence level is high, nod her head [14], smile and have her eyes fully open to give the user feedback that everything is going well and the system is understanding. If, on the other hand, confidence is low, in order to make it clearer to the user that there might be some problem with recognition and that extra care should be taken, an option might be for the ECA to gesture in such a way as to show that she isn't quite sure she's understood but is making an effort to. We have attempted to create this effect by having the ECA lean her head slightly to one side [18], stop smiling and mildly squint. Our goal, once again, is to find out whether these cues do indeed help users realise what the situation is. This is especially important if it helps to avoid the well known problem of falling into error spirals when a recognition error occurs in a spoken dialogue system [24].

In the case of intermediate recognition confidence followed by a mixed initiative strategy involving implicit confirmation, specific gestures could also be envisaged. We have chosen not to include specific gestures for these situations in our first trials, however, so as not to obscure our observations for the high and low confidence cases. A neutral stance for the intermediate confidence level should be a useful reference against which to compare the other two cases.

3.5. Recognition problems

There are basically four types of problems related with recognition. In subsection 3.3 we already talked about no input and phantom input situations. Now we will consider those situations in which the system finds the user's utterance incomprehensible (no-match situations) and those in which the system gets the user's message wrong (recognitions errors)

When a no-match occurs there are two ways in which an ECA can be useful. First, what the character should say must be carefully pondered to ensure that the user is aware that the system didn't understand what she said and that the immediate objective is to solve this particular problem. This knowledge can make the user more patient with the system and tolerate better the unexpected lengthening of the interaction [25]. Second, the ECAs manner should try to keep the user in a positive attitude. A common problem in no-match and error recovery situations is that the user becomes irritated or hyperarticulates in an attempt to make herself understood, which in fact increases the probability of yet another no-match or a recognition error. This we should obviously try to avoid. The ECA behaviour strategy we will test in no-match situations is to have the character lean towards the camera and raise her eyebrows (the idea being to convey a sense of surprise coupled with friendly interest). We have based our gesture on one given in [26]. Figure 1 shows an example dialogue sequence including the association between the different dialogue strategies and the ECA gesture sequences after a user's utterance.

If the user points out to the system that there has been a recognition error in a way that gives the correct information at the same time, then the ECA will confirm the corrected information with special emphasis in speech and gesture. For this purpose we have designed a beat gesture with both hands (see Table 1).

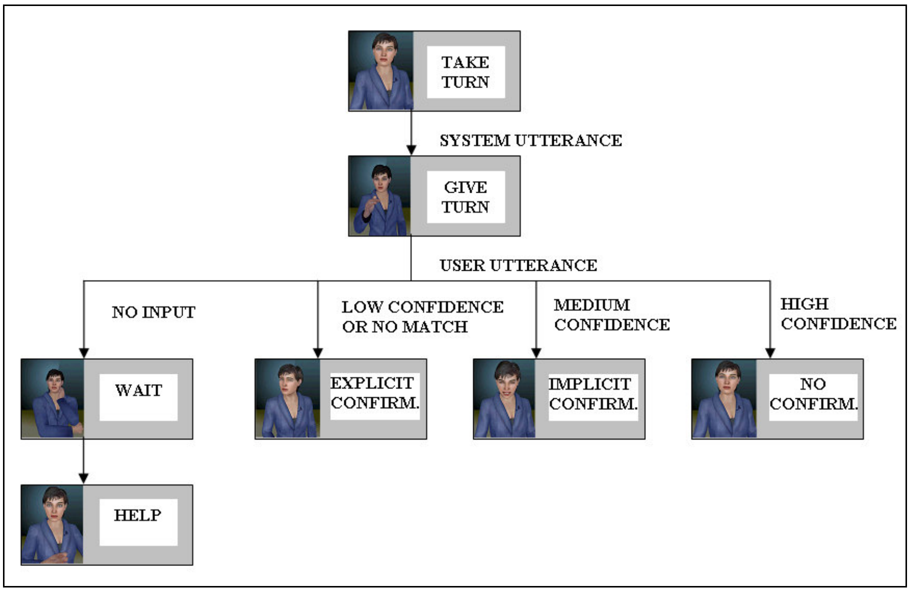

Figure 1. *Dialogue strategies and related gesture sequences.*

3.6. Help offers and requests

It will be interesting to see whether the fact that help is offered by an animated character (the ECA) is regarded by users to be more user-friendly than otherwise. If users feel more comfortable with the ECA, perhaps they will show greater initiative in requesting help from the system; and when it is offered by the system (when a problem situation occurs), the presence of a friendly ECA might help control user frustration.

While the ECA is giving the requested information, she will perform a beat gesture with both hands for emphasis, and she will also change posture. The idea is to see whether this captures the interest of the user, makes her more confident and the experience more pleasant or, on the contrary, it distracts the user and makes help delivery less effective.

Dialogue stage	*ECA behaviour (movements, gestures and other cues)*
Initiation (welcome)	1. Welcoming message: Look at the camera, smile, wave hand 2. Explanation of the task: Zoom in 3. Zoom out, lights dim
Termination	Look forward, nod, wave hand. Turn around, walk away, dim light.
Give turn	Look directly at the user, raise eyebrows, extending open hand towards the user Camera zooms out. Lights dim.
Take turn	Look directly at the user, change posture Camera zooms in. Light gets brighter.
Wait	Slight leaning back, one arm crossed and the other touching the cheek shift of body weight
Help	Beat gesture with the hands. Change posture
Error recovery with correction	Lean towards the camera, beat gesture
Confirmation (high confidence)	Nod, smile, eyes fully open
Confirmation (low confidence)	Slight leaning of the head to one side, stop smiling, mildly squint

Table 1. *Gesture repertoire for the main dialogue stages*

4. EVALUATION ENVIRONMENT

In order to test the various hypotheses put forward in the previous section we have developed an evaluation environment that simulates a realistic mobile videotelephony application (Figure 2) that allows users to remotely check the state (*e.g.*, on/off) of several household devices (lights, heating, etc.).Our dialogue system incorporates mixed initiative, error recovery subdialogues, context-dependent help and the production of guided or flexible dialogues according to the confidence levels of the speech recogniser.

Our environment uses Nuance Communications' speech recognition technology (www.nuance.com). The ECA character has been designed by Haptek, and we have integrated it with our dialogue system, which has required careful planning, designing and programming of all the gesture sequences (each involving a chain of basic movements) using the tools provided with the ECA software (www.haptek.com). (Internally, the ECA is a finite state machine that can be programmed using an editor included in the software.)

Figure 2. *Interface for a mobile device*

Our testing platform allows the collection of the three types of information we well need in the evaluation stage: objective system effectiveness and efficiency parameters (recorded automatically, directly from the users' interaction with the dialogue system, subjective) [27], users' responses to evaluation questionnaires and video recordings of the users interacting with the system.

With all the gathered information we will generate a multimodal evaluation corpus to make the data available to other research groups and to allow comparative analysis between groups. The corpus will contain the *video* recordings of the users' interactions, annotated following standard procedures [28]; annotations of the dialogue states that were generated, following the guidelines in [27]; and *data* comprising the information collected in the questionnaires, which will cover aspects such as frustration, motivation and overall satisfaction [29].

5. TEST DESIGN

We have designed a test plan (subject recruitment, tasks, situations, etc.) with three groups of users: two interacting with the system *with* the ECA and the other without the ECA. For one of the groups with ECA we have programmed the agent with a number of context-specific gestures (i.e., gestures designed to carry or stress information, to capture the user's attention or to make meanings clearer) and a relatively high level of empathic expression (for instance, smiling). The other group will have an ECA with a relatively low level of expression ("neutral" empathy) and which makes no context-specific gestures (for instance, her hands never move). The idea is that the "neutral" ECA act as a reference to compare, on the one hand, with the group without the ECA to see what effects (on performance and user impressions) the sole presence of an ECA might have, and on the other, with the group with the expressive ECA, to study the effects of the different gestures and the overall enhanced expressiveness. The neutral ECA could thus help distinguish between observations related with gestures and expression and those related with ECA presence.

The subjects will carry out the tests in a silent and isolated place so that recognition errors due to ambient noise may be avoided and privacy secured. Before starting the test the subjects will fill out a questionnaire to gather the subjective impressions and expectations of the user. A second subjective questionnaire is presented at the end of the test to see to what extent the users' expectations were met, and how error recovery affected his/her satisfaction with the system. We will also observe users' subjective impression of success (*i.e.*, whether they believe they have carried out the interaction task successfully). In this questionnaire we include videos of the main gestures that the ECA performs at specific points of the dialogue to analyze the users' impressions for each of the gestures, and to find out whether or not they interpret the meaning of the gestures in the intended manner. By means of these questionnaires we hope to be able to validate and, if necessary, modify the nonverbal communication elements that have been introduced throughout the dialogue. (As we have seen, the gestures we have implemented are either adapted from some proposed by leading researchers in the field, or of our own creation. In either case we hope the method of data collection and analysis we have designed will prove useful in defining and refining the gesture repertoire of ECAs in dialogue systems.)

6. CONCLUSIONS AND FUTURE RESEARCH

The absence of studies on real dialogue situations with embodied conversational agents has motivated us to propose a research framework to evaluate the effects of their inclusion in automated dialogue systems, particularly in problem situations.

We have identified a range of problem situations that may arise in dialogue systems, and defined various strategies of ECA use to improve user-machine interaction throughout the whole dialogue.

We have completed the creation of an experimental framework described above to validate, discard or change the proposed strategies, and we are currently in a position to begin testing. The testing sessions will allow us to gather information that will make up a multimodal corpus for analysis and further research.

In future experiments we will attempt to go one step further and analyse how empathic emotions vs. self-oriented behaviour (see [3]) may affect the resolution of a variety of dialogue situations. To this end we plan to design ECA prototypes that incorporate specific emotions, hoping to learn how best to connect empathically with the user, and what effects this may have on dialogue dynamics and the overall user perception of the system.

Conversely, we plan to design tests to study the reactions of users to the emotional behaviour of the ECA, as a first step to modelling different types of users (*e.g.*, extroverted/introverted, patient/irritable, etc.).

ACKNOWLEDGEMENTS

This work has been possible thanks to the support grant received from project TIC2003-09068-C02-02 of the Spanish Plan Nacional de I+D.

REFERENCES

[1] Boyce, S. J., Spoken natural language dialogue systems: user interface issues for the future. In Human Factors and Voice Interactive Systems. D. Gardner-Bonneau Ed. Norwell, Massachusetts, Kluwer Academic Publishers: 37-62, (1999).

[2] Bickmore, T., Cassell, J., Van Kuppevelt, J., Dybkjaer, L. and Bernsen, N. (eds.), *Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems*, chapter Social Dialogue with Embodied Conversational Agents. Kluwer Academic, 2004.

[3] Brave, S, Nass, C, Hutchinson, K, Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int. J. Human-Computer Studies, Nr. 62, Issue 2, pp. 161-178, (2005).

[4] Bell, L. and Gustafson, J., Child and Adult Speaker Adaptation during Error Resolution in a Publicly Available Spoken Dialogue System. Proceedings of Eurospeech 03, Geneve, Schweiz, 2003.

[5] Cassell J., Thorisson K.R., The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence, vol.13, pp.519-538, (1999).

[6] Cassell, J. and Stone, M., Living Hand to Mouth: Psychological Theories about Speech and Gesture in Interactive Dialogue Systems. *Proceedings of the AAAI 1999 Fall Symposium on Psychological Models of Communication in Collaborative Systems*, pp. 34-42. November 5-7, North Falmouth, MA, 1999.

[7] Krauss, R. M., Chen, Y., and Chawla, P. Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (Ed.),

Advances in experimental social psychology (pp. 389-450). San Diego, CA: Academic Press, 1996.

[8] Massaro, D. W., Cohen, M. M., Beskow, J., and Cole, R. A., Developing and evaluating conversational agents. In *Embodied Conversational Agents* MIT Press, Cambridge, MA, 287-318, (2000).

[9] Oviatt, S. & VanGent R., Error resolution during multimodal humancomputer interaction. Proc. International Conference on Spoken Language Processing, 1 204-207, (1996).

[10]Hone, K., Animated Agents to reduce user frustration, in *The 19th British HCI Group Annual Conference*, Edinburgh, UK, 2005.

[11]Oviatt, S., MacEachern, M., and Levow, G., Predicting hyperarticulate speech during human-computer error resolution. Speech Communication, vol.24, 2, 1-23, (1998).

[12] Oviatt, S. Interface techniques for minimizing disfluent input to spoken language systems. In Proc. CHI'94 (pp. 205-210) Boston, ACM Press, 1994.

[13] Poggi, I., How to decide which gesture to make according to our goals and our contextual knowledge. Paper presented at Gesture Workshop 2001 London 18th-20th April, 2001.

[14] Cassell, J., Bickmore, T., Vilhjálmsson, H., and Yan, H., More than just a pretty face: affordances of embodiment. In *Proceedings of the 5th international Conference on intelligent User interfaces*, 2000.

[15]Cassell, J., Nakano, Y. I., Bickmore, T. W., Sidner, C. L., and Rich, C., Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting on Association For Computational Linguistics*, 2001.

[16] Chovil, N. Discourse-Oriented Facial Displays in Conversation. Research on Language and Social Interaction, 25, 163-194, 1992.

[17] Kendon, A. *Conducting interaction: patterns of behavior in focused encounters*, Cambridge University Press, 1990.

[18] San-Segundo, R., Montero, J.M., Ferreiros, J., Córdoba, R., Pardo., J.M., Designing Confirmation Mechanisms and Error Recover Techniques in a Railway Information System for Spanish. SIGDIAL 2001 WORKSHOP. Septiembre 1-2, 2001. Aalborg (Dinamarca).

[19] Oviatt, S., and Adams, B. Designing and evaluating conversational interfaces with animated characters. Embodied conversational agents, MIT Press: 319-345, (2000).

[20] Keeling, K., Beatty, S., McGoldrick, P., and Macaulay, L., *Face Value? Customer Views of Appropriate Formats for Embodied Conversational Agents (ECAs) in Online Retailing*. In Proceedings of the Proceedings of the 37th Annual Hawaii international Conference on System Sciences (Hicss'04), 2004

[21]Schaumburg, H., Computers as tools or as social actors? the users' perspective on anthropomorphic agents. *International Journal of Cooperative Information Systems. 10, 1, 2,* 217-234, (2001).

[22]Catrambone, R, Anthropomorphic agents as a user interface paradigm: Experimental findings and a framework for research. In: *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 166-171), Fairfax, VA, August 2002.

[23] Corradini, A., Mehta, M. and Charfuelan, M., *Interacting with an Animated Conversational Agent in a 3D Graphical Setting*. In: Proceedings of the Workshop on Multimodal Interaction for the Visualization and Exploration of scientific data at the International Conference on Multimodal Interfaces (ICMI'05), October 3-6, Trento (Italy), pp. 63-70, 2005

[24] Bulyko, I., Kirchhoff, K., Ostendorf, M., Goldberg, J., Errorcorrection detection and response generation in a spoken dialogue system. Speech Communication 45, 271-288, (2005).

[25] Goldberg, J., Ostendorf, M., Kirchhoff, K., The impact of response wording in error correction subdialogs, In *EHSD-2003*, 101-106, (2003).

[26] Fagerberg, P., Ståhl, A., Höök, K. (2003). Designing Gestures for Afective Input: an Analysis of Shape, Effort and Valence. In *Proceedings of Mobile Ubiquitious and Multimedia,* MUM 2003, Norrköping, Sweden.

[27] Möller, S., Parameters for quantifying the interaction with spoken dialogue telephone services. ", In *SIGdial6-2005*, 166-177, (2005).

[28] Herramienta de anotación ANVIL: http://www.dfki.de/~kipp/anvil/description.html

[29] Walker, M.A., Litman, D.J., Kamm, C.A. and Abella, A., Evaluating spoken dialogue agents with PARADISE : Two case studies. Computer Speech and Language, 12, 317-347, (1998).