

# ARTIFICIAL CONVERSATIONAL COMPANIONS

## *A Requirements Analysis*

Sviatlana Danilava<sup>1</sup>, Stephan Busemann<sup>2</sup> and Christoph Schommer<sup>1</sup>

<sup>1</sup>*Faculty of Science, Technology and Communication, University of Luxembourg, 6 Rue Coudenhove-Kalergi, Luxembourg*

<sup>2</sup>*Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Stuhlsatzenhausweg 3, Saarbrücken, Germany  
{sviatlana.danilava, christoph.schommer}@uni.lu, stephan.busemann@dfki.de}*

**Keywords:** Artificial companions, conversational agents, human-machine relationship, long-term interaction

**Abstract:** This work is based on several attempts to provide a definition and a design approach of Artificial Companions that can be found in the referenced literature. We focus on computer agents that simulate human language behaviour and are aimed to serve, to assist and to accompany their owner over a long period of time, that we call Artificial Conversational Companions. Although accepted by the research community, the visions set very high expectations of such agents, but they do not address the technical feasibility and the system limitations. This is the first approach to define a set of features that allow an artificial agent to be regarded as an Artificial Conversational Companion. We describe relationships between the components and identify systematic shortcomings of the current systems. We propose a scalable method for implementing the desired capabilities of an Artificial Conversational Companion in a generic framework with reusable, customizable and interdependent components.

## 1 INTRODUCTION

The term *Artificial Companion* (AC) has been introduced in (Wilks, 2006) as “... *an intelligent and helpful cognitive agent which appears to know its owner and their habits, chats to them and diverts them, assists them with simple tasks...*” . The most important characteristics of an AC are the absence of a central task, a sustained discourse over a long time period, a capability to serve interests of the user, and a lot of personal knowledge about the main user (Wilks, 2010).

(Adam et al., 2010) define companions as “... *agents that are intelligent, and built to interact naturally [...] with their user over a prolonged period of time, personalising the interaction to them and developing a relationship with them.*” In (Ståhl et al., 2009) an AC is “*a computational agent that acts as a conversational partner to its user, builds a long-term relationship to the user, and learns about the user's needs and preferences.*” (Webb et al., 2010) emphasise that “*Companions are targeted as persistent, collaborative, conversational partners [which] can have a range of tasks.*”. (Pulman et al., 2010) see a conversation with an AC as “*not necessarily connected to any immediate task*”. (Benyon and Mival, 2008) describe an AC as a “... *personalised conversational,*

*multimodal interface, one that knows its owner.*” They see a companionship as “...*an accessible, pleasing relationship with an interactive source in which there has been placed a social and emotional investment* (Benyon and Mival, 2010).

Summarised, an AC is a personalised, multimodal, helpful, collaborative, conversational, learning, social, emotional, cognitive and persistent computer agent that knows its owner, interacts with the user over a long period of time and builds a relationship to the user.

These visions of an AC raise the level of expectations of such an agent quite high, but they do not address the technical feasibility and the system limitations. Requirements like “to know its owner”, “be helpful” or “long-term relationship” are vague. These requirements and their impacts must be clearly defined in order to implement an AC.

### 1.1 Previous work on companions

(Benyon and Mival, 2010) give an overview on pet and anthropomorphic computer agents. All of them, from Tamagochi to artificial woman, are referred to as “companions”. The form of an AC influences all the issues of interaction and possibilities for companionship (a cat needs only to be a cat, see also (Benyon

and Mival, 2010)). In this paper, we use the term Artificial Conversational Companion (ACC) for companions that are aimed to simulate human language behaviour, in order to distinguish them from those, that are not (e.g. artificial pet companions).

Recent contributions in the domain of ACC are the EU-funded Companions project<sup>1</sup> with the “How Was Your Day” Companion (HWYD-Companion) (Pulman et al., 2010) talking about job-related topics, the Senior Companion (SC) reminiscing about images (Wilks et al., 2011), and the Health and Fitness Companion (HFC) planning daily exercises and diets (Turunen et al., 2011). A pre-companion work on Relational Agents (Bickmore, 2003) focuses on social-emotional relationships between humans and computer agents. The ALIZ-E project<sup>2</sup> concentrates on robot companions for children in a hospital environment (Baxter et al., 2011), and the Child Companion (Adam et al., 2010) is designed to engage a child user with games and stories.

A large amount of research has been done up to now on various questions related to ACCs. However, the implementation of the distinguishing features that are part of the above descriptions of an ACC (personalisation, sustained discourse, required knowledge about the user and the learning mechanism), have not been deeply analysed.

The research results available from the different areas do not yet suffice to fully support an ACC. We will argue that not only domain-specificity but also limited cognitive, social, and emotional competencies will, in spite of the visions put forward, be a factor to be dealt with in the foreseeable future.

## 1.2 Research questions

(Benyon and Mival, 2008) introduce a general model for designing technologies for relationships. The model is based on five concepts: utility, form, emotions, companion’s personality and trust as well as its social attitudes. According to this model, long-term, persistent interactions are part of the companion’s personality and trust axis. However, modelling of what the authors call “companion’s personality” cannot be done independently from modelling its social or emotional attitudes. The sets of the companion’s capabilities shaped alongside these axes will overlap. It will make the model more complex and lead to redundancies.

Moreover, the model does not clearly define what is meant by a “companion’s personality” and how it

should be modelled. The authors mention assertiveness and submissiveness as properties of a personality. However, we cannot observe the properties attributed to a person, but only how the person or agent interacts with another person or agent. This, in turn, depends on how *all* the participants influence the interaction. As will be discussed below, a person interacts differently with different partners – even within the same role distribution like, for instance, in a teacher-student interaction. During long-term interaction, participants tend to adapt their behaviour to the situation and the relationship expectations as well as to the behaviour of the other participants. The usefulness of a notion of an ACC’s personality seems, in our view, dubious.

This paper is the first attempt to define a set of requirements for an ACC that supports the practical implementation. In this paper, we address the following questions:

1. What is the set of mandatory requirements that a computer agent must satisfy in order to be regarded as an ACC?
2. What is, in addition, required for a long-term human-companion interaction?
3. Where are important technical limitations?

We take the view that

- mutual dependencies among the components rather than a system of independent modules will provide the desired functionality;
- the utility and the adaptivity of the system build a basis for a long-term interaction;
- the design of an ACC should accommodate certain future research progress and support different application cases.

We see a possibility to implement such an ACC in a generic system of interdependent components that can be accustomed.

The remainder of this paper is organised as follows: in Section 2 we analyse the requirements and their implementation in the current companion prototypes. We also point out shortcomings and research gaps that may currently make the requirements unattainable. In Section 3 we propose design principles for an ACC framework that takes research gaps and technical limitations into consideration. This is followed by conclusions in Section 4.

## 2 REQUIREMENTS

Which interactional resources are involved in the conversation depends on the modalities involved. For an

---

<sup>1</sup><http://www.companions-project.org/>

<sup>2</sup><http://www.aliz-e.org/>

ACC interacting via instant messages, it will be only text, possibly extended by emoticons etc. For a talking head avatar it will be speech, prosody, eye gaze, facial expressions and head movements. For an embodied agent, the body language and body movements must be added.

The entire interactional phenomena of a talk and the connections among them related to each single participant of an interaction are referred to as *interactional profile* (?). The content, the actions, the flow of emotions and the relationships are determined and co-constructed by all the participants in each interaction and especially in each conversation. We discuss these aspects of the implementation of a companion’s interactional profiles below.

## 2.1 Conversational abilities

Conversation is interactive, spontaneous, exchange of ideas between two or more agents that follow rules of etiquette and politeness, according to social distance and cultural norms. Thus, the conversational part of an ACC is responsible for the understanding and production of spontaneous utterances during the interaction with the user and following rules of social interaction using the available modalities.

Companions are also aimed at maintaining a sustained discourse over a long period of time. This requires an analysis of the conversational history and different cognitive functions to work on it (e.g., associative memory, learning or reasoning). The interaction with an ACC cannot be modelled as just a simple stimulus-response based exchange of utterances.

### 2.1.1 Language understanding and generation

In Natural Language Understanding (NLU), we have the tradeoff between deep language understanding reaching at very elaborate interpretations of utterances at the cost of covering only a restricted domain and shallow language understanding that is unrestricted with respect to the domain but inherently limited in its understanding capabilities – very simple techniques as keyword spotting or pattern matching systematically ignores information available in the text. Work on combining both has been carried out: deep linguistic analysis can be enriched with shallow techniques such as named entity recognition (NER) (Schäfer, 2008). Current ACC prototypes use different shallow techniques for NLU, depending on the system application case. Each of them acts in a single topic domain.

The main objective of the HWYD-Companion was producing longer utterances that are still appropriate in terms of content and emotions. The HWYD-

Companion performs template-based information extraction – it uses shallow syntactic and semantic processing – to find instantiations of event templates. The dialogue manager questions the user until enough slots are filled. Then a longer empathic response is generated.

The HFC uses semantic interpretation for speech recognition and domain specific grammars. “*The Cognitive Manager models the domain [...]. In contrast, the Dialogue Manager focuses on interaction level phenomena, such as confirmations, turn taking, and initiative management.*” (Turunen et al., 2011)

Conversational key features of the SC are reading news from a few categories, telling jokes taken from the Internet, and voice-based picture tagging. The NLU module of the SC is based on GATE (Cunningham et al., 1996). The components have been improved for the SC system by gazetteers containing locations and family relationships. The NER module builds the key part, which is required for this application scenario. The information obtained is then passed to the Dialogue Manager and stored in the knowledge base for later reference.

Understanding the goals underlying the user’s utterances enables a system to decide when to produce the next utterance (turn-taking), what content to convey next, and how to express that content. Complex interactions between dialogue manager, planner and action selection are needed to meet interaction goals like fulfilling some task or just “killing time”.

In classical dialogue applications, templates with a fixed number of slots define how the dialogue should be maintained. Dialogue managers for a free, interactive conversation are mostly improved by enlarging the state space, which leads to combinatorial explosions in planning tasks, as the dialogue develops. This complexity can be managed to a certain extent by policy activation (Kruijff and Lison, 2010). But eventually, all these techniques have the disadvantage that developers need to specify in advance the conditions under which the system produces a certain decision. Similarly, the subsequent system utterance is selected from a set of all probably appropriate utterances. This leads to a perceived repetitiveness of the system, as it has been demonstrated in the evaluation of Laura (Bickmore et al., 2005). Most of the participants found the conversations repetitive at some point. Annoyance and negative feelings reduced the motivation of several participants.

All these systems act in a restricted domain in order to make the NLU task manageable. However, there is no guarantee for the system to always remain within its domain. For instance, in a human-to-human conversation about images, comments like “You look

so beautiful in this picture!” or “Good shot! This dress suits you very well!” are also probable. Moreover, domains are usually not well-defined. For instance, “job-related topics” may include topics on any job, including taxi driver, teacher, programmer, and not only manager.

For an ACC, techniques are required that smoothly open emergency exit doors whenever the system’s coverage is left. Better ways than the infamous “*Tell me more about your family!*” of the ELIZA system are easily available, but depend on available cognitive abilities and social relationships. The straightforward solution – if appropriate – would be to let the user know what the system did not understand.

### 2.1.2 Cognitive abilities

With the vision of an ACC described in Section 1, an ACC needs cognitive functions that allow it to have personal knowledge about its main user, to learn new facts about the user and user’s world, his needs and preferences, and to use this knowledge in the interaction as well. However, cognitive abilities of an ACC do not necessarily have to cover *all* the cognitive functions of humans. It is a subject for the future research to identify the functions required for a personalised long-term interaction, based on the particular application case and desired functionality.

For example, the SC introduced in Section 1.1 performs a real-time information search about places to show that it has some knowledge of what is being talked about while talking about images, in contrast to HWYD-Companion and HFC, that do not have and probably do not need a real time access to additional knowledge and are well served with the facts they have been initialised with.

SC asks the user about the names and the age of the persons on images, and about the places where the pictures were taken. The gathered information is used to maintain the dialogue which is still focused on image tagging. Simple inferences about family relationships it already knows makes the dialogue more natural.

The HWYD-Companion needs for each subtopic a particular amount of information that is extracted from user utterances. Then a longer response of the system is generated.

Two sample interactions show how knowledge about the user in combination with simple inferences improves the quality of the conversation. The first one is by SC: *SC: Where was this photo taken?*

*U: This photo was taken in Tanzania.*

*SC: Have you ever been to Zanzibar?*

*U: Yes, I have.*

*SC: Great place, I hope you saw the beaches. ...*

Then the system goes to the next picture.

These simple inferences (Tanzania, Zanzibar, beaches) make the dialogue more fluent and natural.

In the HWYD-Companion’s video demonstration, the conversation starts usually with *Hello John. How did your work day go today?*, and the user tells the ACC, how it was.

In the demonstration, the user says that he arrived late because of the traffic. The ACC replies: “*You have my sympathy. What happened next?*” A more appropriate reaction would be to ask the user whether he managed it to be on time in the meeting, produced by the inference from the knowledge about the user (the plan to have a meeting) and the context (user arrived late because of the traffic). The given system reaction is emotionally adapted to the context (sympathy) and can be applied in each situation where sympathy would be appropriate reaction.

The existing systems have a predefined information need and just need to get those data from the user that are declared as information need by the programmer. Systems cannot decide, whether or not they need more information as long as there is no deliberate planning associated with the knowledge items mandatory to carry out the plan.

While convincing examples are demonstrated in today’s ACCs, it is completely unclear how the cognitive abilities can be extended to keep the conversation interesting for the user in many-faceted long-term interaction. This requires an amount of inferencing and learning that is not required for the usual task-oriented dialogue systems.

### 2.1.3 Emotional competence

The ability to address the emotional side of companionship may play a key part in their acceptance by the users (Cowie, 2010).

A big progress in affective computing was achieved by the HUMAINE network<sup>3</sup>. Two general types of emotions are studied: pervasive (general personal attitude colouring long time periods) vs. emergent (short, intensive affective states) emotions. In the follow-up project SEMAINE<sup>4</sup> the focus was on non-verbal emotional behaviour and different agent characters (happy, angry or despondent) (Schröder et al., 2011).

Most of this work is not directly applicable to ACCs. Insights on emotion recognition are mandatory but must be complemented by appropriate reactions. Insights on signalling emotions are important,

<sup>3</sup><http://emotion-research.net/>

<sup>4</sup><http://www.semaine-project.eu/>

but need in an ACC trigger methods requiring internal stimuli. Studies of emotional effects in human-computer dialogue are, to our knowledge, not available yet.

Emotion handling in the HWYD-Companion is implemented in form of two feedback loops. The “short loop” provides an immediate backchannel that aligns the companion’s response to the user’s attitude showing, e.g. empathy. The “major loop” is responsible for emphatic utterance generation, typically advice or warning expressed in both verbal and non-verbal behaviour, based on the gathered information.

The SC’s emotional behaviour is based on speech recognition. Recognised emotions were mapped onto a two-dimensional space. SC should be able to recognise user’s emotion placed in this space, formulate a belief about user’s emotional state and move itself in this space for an appropriate response.

Both systems focus only on short and intensive affective states. The recognition of emotions is mainly based on the speech prosody, which is not available in a text based chat. Pervasive emotions need also to be taken into account when designing tools for long-term interaction.

Issues similar to the gathering knowledge about user raise also in emotion handling: the systems can only recognise predefined emotions in particular predefined states, and produce an emotional response in a predefined way. The complexity of emotion handling increases, if issues like “different perception of the same event being in different moods” are considered.

#### 2.1.4 Socio-cultural competence

We communicate as social and cultural entities. There are different sets of rules for successful conversation within different groups of people as this was noticed many decades ago (Knigge, 1805).

The existence of such rules is usually not observable until two legal but contradictory rules are applied by participants of an interaction and lead to a conflict situation and misunderstanding, see for example (Young, 2011; ?).

Currently neither socio-linguistic nor computational models of such interactional rule systems are available, but there are research efforts on socio-linguistic phenomena in discourse (Strzalkowski et al., 2010; Agar, 1996; Scollon and Scollon, 2000). Using small-talk as a form of social dialogue in conversational agents helps to establish a bond between the user and the system (Bickmore et al., 2005). Further research investigations in socio-linguistic phenomena and social signal processing will allow to improve the conversation with an ACC.

The selection and the use of the interactional resources is subject of research in second language acquisition domain. Recent work shows that there is - besides the four common competences: reading, listening, speaking and writing comprehensions - a fifth competence, the *interactional competence*, also referred to as intercultural or transversal (Cook, 1992; Hall et al., 2006). Interactional competence is not fixed, it can be developed in the process of the interaction (Young, 2011).

Each communication has a content part and a relationship part, in which the latter determines the former (Watzlawick et al., 2000). Social and cultural norms and rules are expressed in socio-linguistic phenomena during a conversation. Given this fact, we can eliminate neither the relationship part nor the socio-cultural part from our communication.

Particular implementations of interpersonal relationships between people are in each case unique, but they are categorised in large classes like “friends” or “colleagues”. The relationship between an ACC and its main user will be also belong to one of the large classes that will probably differ from all interpersonal human relationship classes.

Prior to modelling relationship-related speech acts, we need to decide what kind of relationship we want to establish between the user and the ACC. A teacher-learner relationship will necessarily differ from e.g. a personal assistant-boss relationship. The general requirement that an ACC must produce a relational response in its user is vague.

## 2.2 Adaptivity

(Reeves and Nass, 1996) report that users prefer computers that become more like them over time over those which maintain a consistent level of similarity, and that users prefer computers that are similar to them. This property is already used in ALIZ-E, where the robot adapts its behaviour to the user’s behaviour (Baxter et al., 2011).

Several research results show that people adapt their language by selecting the same words and grammar constructions while interacting with other people, but also while interacting with machines, see e.g. (Dobroth et al., 1990). This process is also referred to as convergence and denotes negotiation on vocabulary and communication style among all conversation participants. In this way, artificial agents should be able to adapt their language behaviour to that of their users, who in turn might be influenced by the linguistic behaviour of the ACC.

Since the interaction is co-constructed, interactional profiles (introduced at the beginning of this

section) need to be modelled as a cooperation of the ACC and the user according to a stereotype based user model in the initial implementation. A large amount of personal knowledge is necessary to create a highly adaptive user model of one particular user. The adaptivity mechanism will allow to use the information from the conversations with the user and to adjust the user model.

Besides language behaviour, also cognitive, emotional, and social behaviour is subject to adaptation. While systems that can both interpret and generate such behaviour should in principle be able to adapt themselves to their interlocutor, this requires a way of analysing the (non-)linguistic behaviour and learning why it occurred. A simple mimicking strategy such as using a certain word as frequently as the dialogue partner is bound to fail. We are not aware of relevant studies of credible adaptivity in long-term human-machine interaction, but from ALIZ-E, such results may be expected.

### 2.3 Utility

Tools should be useful. Conversations with an AC require time investment at the expense of time the user could have spent with her family or friends. The services of an ACC compete against services of other machines. There must be a reason why a user decides to use an ACC for searching Internet for news instead of a web browser.

The utility of a companion's services can be taken as a measure of relative satisfaction, which is in this case the frequency of consumed companions services or the cumulative length of the conversations. Experiments with elderly people described in (Benyon and Mival, 2010) show that the kind of services an ACC can perform for its user is considered very important. Desired tasks for a robot companion range from making the tea to doing the ironing.

In contrast to the vision of Wilks, most current AC prototypes have a central task. Bickmore's Laura was aimed to be a fitness coach, the HWYD-Companion is created to talk about job-related topics, and the SC is designed for reminiscing about images.

Not all the user interfaces to computer programs can be replaced by an ACC or a voice control. Corresponding tests have been made with speech-to-text (STT) technology and people who were willing to buy an STT computer, as described in (Savoia, 2011). After a few hours of tests participants changed their mind to the worse. People's throat would get sore, it created a noisy work environment, and it was not suitable for confidential material. However, in scenarios like reading or writing messages, or changing navi-

gation options while driving a car, voice control may well be desirable.

The utility of the system provides a basis for a long-term interaction. Application scenarios need to be elaborated, where an ACC may successfully compete against other devices and conventional computer programs. An ACC could be helpful in applications where the user benefits directly from the conversation, such as conversation training in a foreign language, or where a kind of long-term goal exists as in coaching, teaching, or psychotherapy, when a task-oriented dialogue system will not perform well.

### 2.4 Long-term interaction

Long-term interaction is characterised by the user's continuous motivation to interact with the ACC. Such interaction is influenced by emotions, trust, sympathy, positive emotional bond and/or utility. To keep the user motivated, the interaction must be interesting and stimulating, which can be supported by occasional unexpected utterances or non-linguistic behaviour. Another, more technical requirement is the ACC's need to ensure the consistency and the persistence of the mass data acquired over time in the interactions, and its appropriate usage in, e.g., determining what has been communicated before.

Long-term interactions cannot be enforced. It is based on a continuous interest of the user to communicate. Consider an interaction on a ticket counter, which involves a simple work flow and few dialogue steps to achieve the goal. Taking a few minutes of interaction at most, this is clearly not a suitable ACC application. The requirements differ widely with a person who wants to learn a foreign language and interacts with a teacher in a class at a language school. The teacher needs to keep the person engaged, to respond to personal needs of the learner, and to establish a positive emotional relationship with the learner in order to keep the learner motivated. The teacher needs to acquire much personal knowledge about each learner. While phone ticket selling has been fielded using state-of-the-art dialogue systems, more complex capabilities as described above are necessary for the teaching scenario.

The duration of long-term interaction has always been left open. Certainly it is multi-session interaction, and it extends over several days up to weeks, months, or years. In practice, real-time behaviour of the system will become an issue as mass data accumulate over time, which need to be accessed and interpreted by the ACC. As experiments over extended periods of time have to our knowledge never taken place so far, it is unclear whether a notion of forget-

ting and shaded accessibility of information is mandatory in order to efficiently deal with the data. Results on the functioning of the human brain from further research areas – psycho-linguistics and cognitive science – may be of help here.

### 3 TOWARDS DESIGNING AN ACC SYSTEM

A huge amount of research effort has been invested in different disciplines that are relevant for developing an ACC. Existing ACC systems are not easily comparable due to different design approaches, system architectures and components used. As proposing a full-fledged new ACC architecture would go beyond the scope of this position paper, our notion of “component” is quite vague. Components may correspond to key functionality required for each of the subsections of Section 2. Each subsection may represent multiple components.

Designing an ACC based on the requirements identified above should take the limitations of current systems and the research gaps sketched in the previous sections into consideration. With the current knowledge future changes of ACC system designs and architectures are more than likely. To minimise the frequency with which ACCs must be (re-)built from scratch, we propose a highly modular and customisable approach that will allow components to be defined and reused. Basically this design involves a highly declarative approach to knowledge representation, and – as a consequence – the separation between interpreter components and the knowledge they interpret. This has become a standard approach in NLU, where grammars are kept separate from parsers or generators interpreting them. This way, the knowledge and the interpreters can be modified independently of each other, allowing e.g. to use the same parser for a different language.

The same approach can be applied to modelling mutual dependencies of component behaviour, which – as argued in Section 2.1 – are an integral part of an ACC. First attempts to implement mutual dependencies among the components of an ACC include the NLU and the Cognitive Module in HFC, emotions and language generation in HWYD-Companion, and past history, simple inferences and dialogue management in SC. These systems can already partially integrate the learned information about the user, emotional analysis output, or simple cognitive functions into the conversation.

We propose to generalise on this by virtue of a framework embedding components, component

knowledge, and knowledge about interdependency, through which all interdependent behaviour is modelled in a declarative way. In such a framework, the developer can determine most easily the impact of any change of one element onto any others.

As the conversational competence of ACCs will – in the foreseeable future – be limited to a certain domain and task at a time, particular components will have different complexity for different application cases. For example, the capabilities necessary for a pleasant conversation will be accustomed for a particular service domain. Customisation strongly depends on the modalities of interaction available. For instance, if emotions must be expressed through text, the NLU components must be able to convey appropriate signs. The components will be customised in such a way that they cover the domain and task in hand. We expect that the general architecture by and large remains untouched by customisation, but the knowledge required will very much differ across application cases.

Building various ACCs within this same design paradigm will allow us to compare the particular components of the ACCs and the resulting systems in a unified way.

### 4 CONCLUSIONS

In this paper for the first time a set of features and requirements was proposed that allow an artificial agent to be regarded as an Artificial Conversational Companion. We argue that a multitude of different features must be considered that taken together allow for a useful long-term interaction between an ACC and its user. Since research gaps and technical limitations will prevent the realisation of Wilks’ visionary ideas in the near future, we proposed design principles for the implementation of an ACC framework that is particularly well-suited to extensions in view of scientific and technical progress. Moreover these principles support the customization of the components to dedicated tasks and domains. Mutual dependencies among the components will be explicitly modeled to provide the desired functionality.

### 5 ACKNOWLEDGEMENTS

We would like to thank Prof. Dr. Gudrun Ziegler (University of Luxembourg) for valuable discussions and constructive criticism of an earlier version of this work.

## REFERENCES

- Adam, C., Cavedon, L., and Padgham, L. (2010). "Hello Emily, how are you today?": personalised dialogue in a toy to engage children. In *Proceedings of the 2010 Workshop on Companionable Dialogue Systems*, pages 19–24. ACL.
- Agar, M. H. (1996). *Language Shock: Understanding The Culture Of Conversation*. Harper Paperbacks, 1 edition.
- Baxter, P., Belpaeme, T., Canamero, L., Cosi, P., Demiris, Y., and Enescu, V. (2011). Long-Term Human-Robot Interaction with Young Users. In *IEEE/ACM Human-Robot Interaction 2011 Conf.*
- Benyon, D. and Mival, O. (2008). Landscaping personification technologies: from interactions to relationships. In *CHI '08 extended abstracts on Human factors in computing systems*, CHI EA '08, pages 3657–3662.
- Benyon, D. and Mival, O. (2010). From human-computer interactions to human-companion relationships. In *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia*, IITM '10, pages 1–9.
- Bickmore, T. W. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. PhD thesis, Massachusetts Institute of Technology.
- Bickmore, T. W., Rosalind, and Picard, W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer Human Interaction*, 12:293–327.
- Cook, V. (1992). Evidence for multi-competence. *Language Learning*, 42(4):557 – 591.
- Cowie, R. (2010). Companionship is an emotional business. In *Close Engagements With Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. John Benjamins Publishing Company.
- Cunningham, H., Wilks, Y., and Gaizauskas, R. (1996). Gate – a general architecture for text engineering. In *In Proceedings of COLING-96*. ACL.
- Dobroth, K., Karis, D., and Zeigler, B. (1990). The design of conversationally capable automated systems. In *Proceedings of the 13th International Symposium on Human Factors in Telecommunications*, pages 389–396.
- Hall, J. K., Cheng, A., and Carlson, M. T. (2006). Reconceptualizing multicompetence as a theory of language knowledge. *Applied Linguistics*, 27(2):220–240.
- Knigge, A. F. (1805). *Practical philosophy of social life: or, The art of conversing with men: after the German of Baron Knigge*. Penniman & Bliss, O. Penniman, printers, Troy.
- Kruijff, G. J. and Lison, P. (2010). Policy activation for open-ended dialogue management. In *Proceedings of the AAAI 2010 Fall Symposium “Dialogue with Robots”*. AAAI.
- Pulman, S. G., Boye, J., Cavazza, M., Smith, C., and de la Cámara, R. S. (2010). 'how was your day?'. In *Proceedings of the 2010 Workshop on Companionable Dialogue Systems*, pages 37–42. ACL.
- Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places (CSLI - Lecture Notes)*. Cambridge University Press.
- Savoia, A. (2011). Pretotype it. [googletesting.blogspot.com/2011/08/prototyping-different-type-of-testing.html](http://googletesting.blogspot.com/2011/08/prototyping-different-type-of-testing.html).
- Schäfer, U. (2008). Shallow, deep and hybrid processing with uima and heart of gold. In *Proc. of the LREC-2008 Workshop Towards Enhanced Interoperability for Large HLT Systems: UIMA for NLP*, 6th Int. Conf. on Language Resources and Evaluation, pages 43–50.
- Schröder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., ter Maat, M., McKeown, G., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., de Sevin, E., Valstar, M., and Wöllmer, M. (2011). Building autonomous sensitive artificial listeners. *IEEE Transactions on Affective Computing*, 99(1).
- Scollon, R. and Scollon, S. W. (2000). *Intercultural Communication: A Discourse Approach (Language in Society)*. Wiley-Blackwell, 2 edition.
- Ståhl, O., Gambäck, B., Turunen, M., and Hakulinen, J. (2009). A mobile health and fitness companion demonstrator. In *EACL '09*, pages 65–68.
- Strzalkowski, T., Broadwell, G. A., Stromer-Galley, J., Shaikh, S., Taylor, S., and Webb, N. (2010). Modeling socio-cultural phenomena in discourse. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1038–1046. ACL.
- Turunen, M., Hakulinen, J., Ståhl, O., Gambäck, B., Hansen, P., Rodríguez Gancedo, M. C., de la Cámara, R. S., Smith, C., Charlton, D., and Cavazza, M. (2011). Multimodal and mobile conversational health and fitness companions. *Comput. Speech Lang.*, 25:192–209.
- Watzlawick, P., Beavin, J. H., and Jackson, D. D. (2000). *Menschliche Kommunikation. Formen, Strungen, Paradoxien*. Huber, Bern.
- Webb, N., Benyon, D., Hansen, P., and Mival, O. (2010). Evaluating human-machine conversation for appropriateness. In *Proceedings of LREC2010*.
- Wilks, Y. (2006). Artificial companions as a new kind of interface to the future internet. Technical report, Oxford Internet Institute.
- Wilks, Y. (2010). Is a companion a distinctive kind of relationship with a machine? In *Proceedings of the 2010 Workshop on Companionable Dialogue Systems*, pages 13 – 18. ACL.
- Wilks, Y., Catizone, R., Worgan, S., Dingli, A., Moore, R., Field, D., and Cheng, W. (2011). A prototype for a conversational companion for reminiscing about images. *Computer Speech & Language*, 25(2):140 – 157.

Young, R. F. (2011). Interactional competence in language learning, teaching, and testing. In Hinkel, E., editor, *Handbook of research in second language teaching and learning*, volume 2, pages 426–443. London & New York: Routledge.