ARTIFICIAL CONVERSATIONAL COMPANIONS
Requirement analysis

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

Abstract: This paper is based on several attempts to provide a definition of an Artificial Companion that can be found in the referenced literature. Although accepted by the research community, such descriptions set very high expectations of such agents, but they do not address the technical feasibility and the system limitations, that is why they need an elaborated and precise explanation. In this paper, 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. This is the first approach to analyse the effort required to meet the expectations of Artificial Conversational Companion and discuss the identified design issues.

1 INTRODUCTION

The term *Artificial Companion* (AC) has been introduced by Y.Wilks 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..."* (Wilks, 2006). The most important characteristics of an AC are absence of a central task, a "never-ending" conversation, sustained discourse over a long time period, a capability to serve interests of the main user and a lot of personal knowledge about the main user (Wilks, 2010b).

(Benyon and Mival, 2008) describe an AC as *"... personalised conversational, multimodal interface, one that knows its owner."* In (Benyon and Mival, 2010) they comment that *"companionship is about an accessible, pleasing relationship with an interactive source in which there has been placed a social and emotional investment. There is a level of trust, compatibility and familiarity within this relationship that results in a feeling of security, content and general wellbeing."* (Adam et al., 2010) define Companions as *"... agents that are intelligent, and built to interact naturally (via speech and other modalities) 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 users 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) describes a conversation with an AC as "not necessary connected to any immediate task".

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 (long-term) relationship to the user. An AC should simulate a human companion in terms of "*one paid to accompany or assist or live with another, one employed to live with and serve another*" (Merriam-Webster, 2009).

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 and must be clearly defined in order to build 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-Companions project[1] with the implementation of a "How Was Your Day" Companion (HWYD-Companion) (Pulman et al., 2010) for talking about job-related topics with the user, the Senior Companion (SC) for reminiscing about images (Wilks et al., 2011), and the Health and Fitness Companion (HFC) for daily exercise planning, leisure activities and diet (Turunen et al., 2011). T.Bickmore's pre-companion work on Relational Agents with the implementation of a personal coach Laura (Bickmore, 2003) focuses on social-emotional relationships between humans and computer agents. The ALIZ-E project[2] concentrates on robot Companions for children with metabolic diseases 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. There are also several investigations in an early stage. Key social, psychological, ethical and design issues are discussed in (Wilks, 2010a).

There is a large amount of research work done to date concerning different questions related to ACCs. Often discussed questions are: emotion, politeness and affection (Cowie, 2010), appearance (Romano, 2010), communication modes, the number of Companion's personalities, ethics and moral (Vargas et al., 2011; Bryson, 2010), trust (Bickmore, 2003, pp.38-40), data protection and privacy (Boden, 2010), goals, world knowledge. However, there is no discussion in the literature concerning the distinguishing features of ACCs that are part of the above descriptions of ACC, such as implementation of a relationship, sustained discourse, defining required knowledge bout the user and describing the learning mechanism, the minimum length of a long-term period and so on.

1.2 Research questions

(Benyon and Mival, 2008) introduce a general model for designing technologies for relationships. Their star model of designing for relationships is based on five concepts: utility, form, emotions, Companion's personality and trust and its social attitudes as well. According to this model, long-term, persistent interactions are part of Companion's personality and trust axis. However, cognitive, emotional and socio-cultural properties form Companion's personality, and its ability to learn and to adapt its behaviour (in all senses) to user's behaviour build a basis for long-term support of user's interest, but utility is also crucial for long-term interaction. Therefore this star model is not minimal and the axes are not disjunct.

[1] http://www.companions-project.org/

[2] http://www.aliz-e.org/

This paper is our first attempt to define a set of requirements for an ACC - a computer agent that is able to serve, to assist and to accompany its main user, and to simulate human intelligence as well. In this paper, we address the following questions:

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

We take the view that mutual dependencies among the components rather than a system of independent modules will provide the desired functionality, that utility of the system builds a basis for a long-term interaction, that the complexity of particular components can be adjusted dependent on the application case. We see a possibility to implement such an ACC in a generic system of interdependent components that can be parametrised.

The rest of this paper is organised as follows: in Section 2 we analyse the requirements and their implementation in the current companion prototypes. In Section 3 we discuss the technical limitations. This is followed by conclusions in Section 4.

2 REQUIREMENTS

(Webb et al., 2010) emphasise that good companions need to be good conversational partners.

In contrast to our conventional devices, the interaction with a Companion should not necessary be connected to any immediate task. Interaction goals like "killing time" or "accompanying some manual work by gossip quite unconnected with what they are doing" (Malinowski, 2004) are examples for such situations. To satisfy this requirement, ACC needs to be able to maintain a conversation with the user.

2.1 Conversational abilities

Conversation is interactive, spontaneous, exchange of ideas by spoken words between two or more persons that follow rules of etiquette, 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 by spoken or written words (chat, instant messages), depending on interaction modalities involved, and following rules of social interaction.

Companions are aimed to maintain a sustained discourse over long period of time, so they need to be more sophisticated conversational partners than chatbots. This requires different cognitive functions to be involved in the interaction, such as a kind of associative memory, learning, reasoning, understanding user's emotional state and appropriate emotional response. Thus, interaction with an ACC cannot be modelled as just a simple stimulus-response based exchange of utterances. A more complex model for conversation is required.

2.1.1 Cognitive abilities

An ACC must have selected cognitive capabilities to successfully serve, assist and accompany its main user. With the above vision of an ACC as a start point, an ACC needs cognitive functions that allow it to have personal knowledge about its main user, learn about users needs and preferences, learn new facts about the user and user's world, use this knowledge in the interaction with the user.

However, cognitive abilities of an ACC do not necessarily have to cover all the cognitive functions of humans. It is rather interesting to identify the functions required for a personalised long-term interaction. The selection of the required cognitive abilities needs to be done based on the particular application case and desired functionality. E.g SC (see Sec. 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.

Current ACC prototypes ask the user explicitly about personal information, preferences, daily activities, environment etc. SC asks the user about the names and 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 from family relationships it already knows makes the dialogue more natural. SC "retains all user input for reference in repeat user sessions, in addition to the knowledge base that has been updated by the Dialogue Manager on the basis of what was said." A hybrid-world approach is used as a world knowledge model.

The HWYD-Companion needs for each subtopic a particular amount of information, that is extracted from users utterances. This information need will drive the dialogue manager to ask the user questions until the required amount of information is available. Then a longer response of the system is generated.

How knowledge about the user combined with simple inferences influences the conversation quality, can be shown by comparing two demonstrations, one of SC and one of HWYD-Companion:
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. (aber diemeisten menschen würden kommentare wie "du siehst hier jung aus" oder "schönes bild, gute komposition" oder "so was hat man damal getragen?!" abgeben, und nicht zu jedem bild "give me the next picture, where was this picture taken")

In the HWYD-Companion's video demonstration, the AC is aimed to chat with the user daily about job related topics, 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?*" To know the user means in this context to have and to retrieve the information about user's plans for that morning (the user planned to have a meeting) and to infer from that plan and the context (user arrived late because of the traffic) that the user may have missed the meeting. A more appropriate reaction would be to ask the user whether he managed it to be in the meeting on time. The actual system reaction is emotionally adapted to the context (sympathy), but the content is very impersonal and can be applied in each situation where sympathy would be appropriate reaction.

2.1.2 Emotional competence

(Picard, 1995) highlights several results from the neurological literature which indicate that emotions play a necessary role not only in human creativity and intelligence, but also in rational human thinking and decision-making. Based on this evidence, to build Companions that maintain a smart conversation, make intelligent decisions and are collaborative requires building Companions that simulate emotions.

(Cowie, 2010) mentions that ability to address the emotional side of companionship may play a key part in acceptance. He warns engineers of inserting Companions into emotionally sensitive roles without engineering them to take that into account.

Emotion handling in the HWYD-Companion is implemented in form of two feedback loops: the short

loop (response time below 700 ms) and the major loop (response time below 3000 ms). The former provides an immediate backchannel, aligns the companions response to the users attitude showing empathy. The latter 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. Mapping a range of emotions onto a two dimensional space was selected as a representation. Reinforcement learning has been used to train the system for emotional utterances. Recognised emotions are 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. To handle emotions means means also to recognise user's emotional type of personality (optimistic, aggressive etc.) and use this information in shaping the communication.

A big progress in affective computing was achieved by HUMAINE project[3] Two general types of emotions are studied: pervasive (general personal attitude) vs. emergent (short, intensive affective states) emotions. This was , followed by SEMAINE project[4] with the focus on non-verbal emotional behaviour and types of agent's personality (Schröder et al., 2011).

2.1.3 Socio-cultural competence

There are different sets of rules for successful conversation within different groups of people, this was notices many decades ago (Knigge, 1805). This means to meet social and cultural norms and rules that are expressed in socio-linguistic phenomena, that is why *socio-cultural competence* is an important aspect of the implementation of an ACC.

There is a set of shared rules for each language group and for each culture that help to avoid unpleasant moments in a conversation and to reach the communication goals. 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. A good example of such situations is described in (Young, 2011), where a Navajo (the father of a boy from the English class) communicates with an English teacher, both are polite in each language culture, but Navajos communication style seems to be impolite for the English teacher and vice versa.

So in this case both communication parts have an intention to adapt their language to parent-teacher-situation, however they select wrong - in terms of socio-cultural context - interactional resources. An interaction between human user and an ACC is co-constructed, too. Thus, an ACC should be able not only to understand and generate grammatically correct sentences, but also share a set of social and cultural rules with the user, and anticipate the user's state.

Sharing the same social and cultural rules means feeling group belonging, which is an important social tie. The state-of-the-art Companions realise "accompany" as interaction with the user, mostly talking about topics that may be relevant for the particular user group. The Child Companion should be also able to play games and tell stories, because it is specific for this user group.

To date, there are neither socio-linguistic nor computational models of such interactional rule systems, 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.

2.1.4 Natural language understanding

In the Natural Language Understanding (NLU) domain, we still have the tradeoff between a deep language understanding in a restricted domain and shallow language understanding based on such simple techniques as keyword spotting or pattern matching within a wide range of topics. There is also a possibility to enrich a shallow understanding with e.g. named entity recognition. Current ACC prototypes use different 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, by e.g giving the user an advice, and simulating conversational scenarios where people tell longer stories without being asked to do it explicitly. The HWYD-Companion performs a template-based information extraction using shallow syntactic and semantic processing to find instantiations of event templates. The dialogue manager questions the user until enough slots are filled.

[3] http://emotion-research.net/

[4] http://www.semaine-project.eu/

Then a longer emphatic response is generated.

HFC uses semantic interpretation for speech recognition and domain specific grammars. "Dialogue management is based on close- cooperation of the Dialogue Manager and the Cognitive Manager. The Cognitive Manager models the domain, i.e., knows what to recommend to the user, what to ask from the user, and what kind of feedback to provide on domain level issues. 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 Senior Companion are reading news from a few categories, telling jokes taken from the internet, and a voice based picture tagging. The ability to make simple inferences helps to produce short sub-dialogues that are not directly connected to the picture tagging task. The NLU module of the SC is based on GATE (Cunningham et al., 1996). Tokenizer, sentence splitter, POS tagger, parser and Named Entity Recogniser (NER) are involved in the process. NER module builds the key part, which is required for this application scenario. These components have been improved for the SC system by gazetteers containing locations and family relationships. The information obtained is then passed to the Dialogue Manager and then stored in the knowledge base.

2.2 Adaptivity

Citing Kelley, Bickmore and Picard say that relationships demonstrate interdependence between two parties a change in one results in a change to the other (Bickmore et al., 2005).

(Reeves and Nass, 1996) report in their book on Media Equation 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 match them in personality. Adaptivity principle allows ACC to become similar to user's personality even if the personalities are quite different at the start time. 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. This seems to be a necessary condition for a comfortable interaction, thus an ACC must be able to do the same. In this way, artificial agents and human users influence each others language behaviour, so the interdependency of the behaviour provides the necessary conditions for a relationship.

Using similar language means having similar personalities. The adaptivity mechanism will allow the system become more similar to the user over time.

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. If we want to build an AC for people with a "normal" social life, it must be useful, helpful and engaging. If there are also other kind of services (assist with simple tasks, be helpful) that an ACC is able to do, it competes against other machines and/or computer programs. There must be a reason why a human user decides to use an ACC for searching Internet for news instead of a web browser. We will use the term *utility* for the set of Companion's services, and the usability of those.

The utility of Companion's services could 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 conversation. Experiments with elderly people described in (Benyon and Mival, 2010) show that it it a very important question, what an AC can do for the users, how can an AC serve its owner. E.g. desired tasks for a robot companion range from making the tea to doing the ironing.

Thus the goal is to build a system that is useful but does not have a central task according to the vision of Wilks. However, the most of existing AC prototypes do have a central task, which is also necessary due to domain restriction. E.g. Bickmore's Laura was aimed to be a fitness trainer, HWYD-Companion is created to talk about job related topics, and the SC is designed for reminiscing about images or get some news from the Internet.

Having a central task does not necessary mean being helpful or able to serve and assist the user. A concept of companion driven picture tagging was studied in (Benyon and Mival, 2010) in a Wizard-of-Oz experiments with 40 people. The main finding reported is that people get bored quite quickly.

Not all the interfaces to computer programs can be replaced by the ACC or a voice interface. Corresponding tests have been made e.g. with speech-to-text (STT) technology. A Wizard-of-Oz experiment with people who were willing to buy a STT computer are described in(Savoia, 2011). After a few hours of

tests participants changed their mind. Peoples 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 while driving a car it may be desirable to have a voice control. An ACC could assist in planning the route or changing navigation options.

2.4 Long-term interaction

We have here a twofold situation: (1) the user's motivation to interact with the ACC that is influences by emotions, trust, sympathy, positive emotional bond or utility; and (2) the ACC needs to guarantee the consistence, persistence of the data stored, the storage needs to be large enough and reliable (data security, privacy etc.), it might be necessary to migrate Companion's "mind" to different hardware because hardware technology develops very quickly. What will happen to an out-of-date robot companion? How can a life-long companionship be guaranteed? Companion designers will face these and many other issues.

Concerning the long-term interaction we need to determine first of all, how long the minimum length of "a long period of time" is. There are different hardware and software requirements (e.g memory size, architecture, data representation, data structures etc.) for an agent that should communicate with one user several weeks long, than for an agent, that is aimed to interact with one user for decades. In the context of an interaction with a computer agent, long-term interaction means open-end interaction without any predefined end point. We use the term *open-end* for an interaction that does not have any predetermined conditions for termination.

3 SYSTEM LIMITATIONS

There are first attempts to implement mutual dependencies among the components of an ACC (NLU and Cognitive Module in HFC, emotions and language generation in HWYD-Companion, past history, simple inferences and dialogue management in SC)

These systems already partially integrate new "learned" information about the user or emotional analysis output or simple cognitive functions into conversation. For a successful communication, *all* these components and a few components that have not been considered in these systems, are necessary.

Even in a simplified scenario of verbal communication with an ideal level of emotion handling, the system needs to understand user's utterances and to decide *when* to produce the next utterance (turn-taking) and *what* to say next (and then *how* to express the selected meaning). It is a non-trivial task for dialogue manager, planner and action selection to meet the interaction goals like "killing time" or "accompanying some manual work by gossip". In classical dialogue applications, templates with a strong, predefined 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 the planner of an exponential complexity, and makes the dialogue developing to a non-trivial task. This high complexity can be slightly managed e.g. by policy activation mechanism (Kruijff and Lison, 2010). But finally, all these techniques have the disadvantage that the programmer needs to know in advance, under which conditions the system is allowed to fire. The system can make a decision only in particular predefined situations, even if it is fine-grained. Similarly, the decision, what the next utterance of the system will be, means selection from a set of all probably appropriate utterances. This leads to a perceived repetitiveness of the system.

Similar issues raise in gathering knowledge about user or emotion handling: the existing systems have predefined information need, they just need to get the infos from the user, that are declared as information need by the programmer. Systems cannot decide, whether they need more information or not, if they do not have a kind of template with empty slots or questionnaire to each topic . They 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 in the design of the emotional competence of an ACC.

Concerning Companion's personality, the decisions of the components are based on statistic data (whether machine learning techniques or finite state automata), however each of the data producers does not behave like the average.

There is no unique measure for conversation quality, and there is a large amount of work done on evaluating spoken dialogue systems e.g., (Danieli and Gerbino, 1995; Artstein et al., 2008). (Webb et al., 2010) suggest appropriateness as a measure of conversation quality, especially for ACs, where human annotators score the level of appropriateness.

Since conversation is a social phenomenon, the feasibility of maintaining a relationship with the user and the quality of conversation cannot be separated, which is well demonstrated in the evaluation of Laura.

(Bickmore et al., 2005) report that most of the participants found the conversations repetitive at some point during the month. This repetitiveness annoyed the participants, and a few of them even indicated that it negatively influenced their motivation.

Particular implementations of interpersonal relationships between people are in each case unique, but they are categorised in large classes like "friends", "colleagues" or "enemies". The relationship between an ACC and its main user will be unique, but it will belong to a large class. Each communication has a content part and a relationship part, in which the latter determines the former (Watzlawick et al., 2000). This statement is called *metacommunicative axiom*. Given this fact, we cannot eliminate the relationship part from our communication, thus it is not required to construct a "relational language" like it is done in most of investigations on relational agents, see e.g. (Bickmore et al., 2005). Each verbal and non-verbal expression contains our relationship expectations and intentions. What exactly we wrap into our words and gestures, depends on the kind of relationship we expect to have with our conversational partners. We signalise social distance or closeness to our conversational partners, but even signalising a large distance means signalising a relationship intention. Prior to start modelling relationship-related speech acts, we need to decide, what kind of social-emotional relationship we want to achieve between the user and the ACC. The general requirement that an ACC must produce a relational response in its user is ambiguous, and is even covered by the composition of conversation, cognitive capabilities, emotional and sociocultural competences and adaptivity.

A multimodal interaction is not an urgent requirement for companionship, sometimes it is even annoying as (Bickmore et al., 2005) observed. For real long-term human relationships it is often acceptable to communicate via text messages, real-life situations are letter friendships or long-distance relationships. Teenagers or young adults will be more likely to communicate with other people only via email or instant messenger or postings in social networks than seniors or children, so they are more likely to accept an ACC which is only able to write short text messages e.g. using instant messenger (IM). In this case ACC would be one of the contacts in the contact list, the chat situation will be more natural for a human user familiar with an IM. For this reason we see multimodal interaction as an optional requirement.

The capabilities of an ACC can be implemented in form of a generic parametrisable system with reusable components. Particular components will have different complexity for different application cases. E.g. capabilities necessary for a pleasant conversation will be parametrised for a particular service domain.

High interdependency of the component will not allow to parametrize all the components independently, but general configurations for similar application scenarios can be performed.

4 CONCLUSIONS

TBD morgen

Es gibt in jedem derGebiete interessante Entwicklungen, man kann davon profitieren, wenn man diese zusammenbringt.

Generisches parametrisierbares system.

Mutual dependency of the components rather than a system of independent modules.

Problems that raise with this system model: Since the complexity of the components depends on the application, the parametrisation cannot be done

Meine Position in 5 Zeilen: akzeptieren, dass diese Vision so nicht erfüllbar ist. Anwendungsfälle für ACC-Systeme finden und die Systeme je nach Anwendung parametrisieren. Ähnliche technische Randbedingungen, aber unterschiedliche Aufgaben.

Given the requirement to accompany the user, thus, to have a social function, social and cultural aspects of the human interaction need to be modelled and integrated into the conversation with Companions.

Utility of Companion's services is one of the basic condition for an open-end interaction. We need to identify scenarios in which an ACC successfully competes against other devices and conventional computer programs.

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.

Artstein, R., Gandhe, S., Leuski, A., and Traum, D. (2008). Field testing of an interactive question-answering characte. In *ELRA Workshop on Evaluation, at ELRA 2008*, pages 36–40.

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 Conference (Robots with Children Workshop)*.

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.

Boden, M. A. (2010). Conversationalists and confidants. In *Close Engagements With Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. John Benjamins Publishing Company.

Bryson, J. J. (2010). Robots should be slaves. In *Close Engagements With Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. John Benjamins Publishing Company.

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.

Danieli, M. and Gerbino, E. (1995). Metrics for evaluating dialogue strategies in a spoken language system. In *Proceedings of the 1995 AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*.

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.

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.

Malinowski, B. (2004). *Magic, Science and Religion and Other Essays 1948*. Kessinger Pub Co.

Merriam-Webster (2009). *Merriam-Webster's Collegiate Dictionary, 11th Edition*. Merriam-Webster, Inc.

Picard, R. W. (1995). Affective computing. Technical report, MIT.

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.

Romano, D. M. (2010). Look, emotion, language and behavior in a believable virtual companion. In *Close Engagements With Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. John Benjamins Publishing Company.

Savoia, A. (2011). Pretotype it. googletesting.-blogspot.com/2011/08/pretotyping-different-type-of-testing.html.

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.

Vargas, P., Fernaeus, Y., Lim, M., Enz, S., Ho, W., Jacobsson, M., and Ayllet, R. (2011). Advocating an ethical memory model for artificial companions from a human-centred perspective. *AI & Society*, pages 1–9.

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., editor (2010a). *Close Engagements With Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. John Benjamins Publishing Company.

Wilks, Y. (2010b). 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.