

## A Configurable Conversational Agent to Trigger Students' Productive Dialogue: A Pilot Study in the CALL Domain

Stergios Tegos · Stavros Demetriadis · Thrasyvoulos Tsiatsos

Published online: 22 October 2013

© International Artificial Intelligence in Education Society 2013

**Abstract** Conversational agents constitute a specific type of ITSs that has been reportedly proven successful in helping students in one-to-one settings, while recently their impact has also been explored in computer-supported collaborative learning (CSCL). In this work, we present MentorChat, a dialogue-based system that employs a configurable and domain-independent conversational agent for triggering students' productive dialogue. After a system overview with an emphasis on design rationale and system architecture, we present a pilot study, where the agent is evaluated in the context of the computer-assisted language learning (CALL) domain. Thirty students collaborated in small groups trying to accomplish three different tasks. MentorChat conversational agent supported each group discussion differently in each task providing (a) 'weak'-directed interventions and/or (b) undirected interventions. The study results indicate that 'weak'-directed agent interventions can be more effective than undirected interventions by means of increasing the level of explicit reasoning, and thus productive dialogue. Some encouraging results concerning the system usability and user acceptance are also presented.

**Keywords** Intelligent tutoring systems · Conversational agents · Computer-supported collaborative learning · Computer-assisted language learning

---

S. Tegos (✉) · S. Demetriadis · T. Tsiatsos  
Informatics Department, Aristotle University of Thessaloniki, Thessaloniki, Greece  
e-mail: stegos@csd.auth.gr

S. Demetriadis  
e-mail: sdemetri@csd.auth.gr

T. Tsiatsos  
e-mail: tsiatsos@csd.auth.gr

## Introduction

The development of Intelligent Tutoring Systems (ITS) has already had a long tradition in technology-enhanced learning (Sleeman and Brown 1982). The evolution of ITS along with research in cognitive psychology and artificial intelligence gave rise to the concept of the “pedagogical agent” (Wenger 2004). In computer science, a pedagogical agent can be regarded as an entity that has specific learning goals and ‘acts’ autonomously in educational settings (Gulz et al. 2011).

Considering the conversational capabilities of recent ITSs, a new subgroup of pedagogical agents has emerged: “conversational agents” (Kerly et al. 2009). Conversational agents, typically, engage learners in a conversation (or simple communication) by using natural language, body movements or even facial expressions (Gulz et al. 2011).

Our work focuses on adaptable conversational agents that use natural language to scaffold peer discussions in collaborative learning settings. The use of conversational agents in the context of Computer-Supported Collaborative Learning (CSCL) constitutes a relatively new area of research aiming to improve students’ learning experience and outcomes (e.g. Chaudhuri et al. 2008; Kumar and Rosé 2011).

In the following, we present: a) the theoretical background of our work, b) the overview of a prototype dialogue-based ITS, namely MentorChat, which uses a conversational agent to promote students’ productive dialogue, and c) a pilot study conducted to contrast the impact of two different agent intervention strategies. The pilot study results provide encouraging evidence regarding the user acceptance of the system.

## Theoretical Background

### Conversational Agents

A conversational agent can hold a discussion with students in different ways, and the type of conversation may vary from spoken (Wik and Hjalmarsson 2009) to text-based (Chaudhuri et al. 2009) or even non-verbal (Ruttkey and Pelachaud 2004). Likewise, the visual appearance of the agent may be human-like or cartoonish, static or animated, two-dimensional or three-dimensional (Dehn and van Mulken 2000). Conversational agents that have a visual representation are often labeled as “embodied” or “animated” conversational agents (Cassell et al. 2000), especially in situations where it is important for the agent to appear more lifelike. In fact, it was found that the visual appearance and ‘personal’ features are crucial to designing an effective pedagogical agent, regardless of the level of its computational functionality (Baylor 2009; Karacora et al. 2012). Over the past decade, conversational agents have been developed to serve multiple pedagogical roles, such as tutors, coaches or learning companions (Haake and Gulz 2009). In addition, many dialogue-based ITSs have utilized conversational agents to meet a variety of educational needs such as question-answering (Feng et al. 2006), tutoring (Heffernan and Croteau 2004; VanLehn et al. 2007), language learning practice (Hjalmarsson et al. 2007; Wik and Hjalmarsson 2009), and the promotion of health-related behavioral changes (Kennedy et al. 2012) and of metacognitive skills (Kerly et al. 2008).

Many of the conversational agents developed in the past aimed at engaging students in one-to-one (student-agent) tutorial dialogues (Evens and Michael 2006; Graesser et al. 2008; VanLehn et al. 2007). Such agents typically simulated the instructional behavior of a human tutor by exploiting AI techniques to converse with students on pre-determined topics. Based on the observations of VanLehn et al. (2007), a student-agent tutorial dialogue can offer several benefits over a monologue, such as the detection and remediation of failed communication, the correction of inaccurate student knowledge and increased interactivity.

Indeed, a lot of successful stories (e.g. Why2-Atlas, AutoTutor, ITSPOKE, Ms Lindquist, CALMsystem, ReTuDiS) were reported regarding the beneficial use of such agents in individual learning settings (Alevan et al. 2004; Graesser et al. 2005a; Grigoriadou et al. 2005; Heffernan and Croteau 2004; Litman and Silliman 2004; VanLehn et al. 2007). Among other findings, the results indicated learning and memory gains (Graesser et al. 2005b; VanLehn et al. 2007), an increase in student motivation (Heffernan and Croteau 2004) and improvement of self-explanation (Alevan et al. 2004; Grigoriadou et al. 2005) as well as of self-assessment skills (Grigoriadou et al. 2005; Kerly et al. 2008).

Nevertheless, a literature review reveals some points of criticism regarding the impact of conversational agents on individual learning, as follows:

- The shallow interactions often taking place between the student and the agent (Jia 2004). Rosé et al. (2004) found that students frequently provided more simplified answers to the conversational agent (e.g. one-word responses) as compared to their answers to the human tutors;
- The agents' role as teacher substitutes often leading to the exclusion of teachers from the educational process (Shawar and Atwell 2007);
- The agents' strict dependence on the instructional domain that may limit (a) the agent's ability to adapt and operate in a new learning context (Dyke et al. 2012) and (b) the teacher's ability of configuration (Kinshuk 2002).

### Conversational Agents in CSCL

More recently, researchers turned their attention to designing ITS systems utilizing conversational agents for supporting collaborative learning activities (e.g. Chaudhuri et al. 2009; Dyke et al. 2012; Kumar and Rosé 2011; Stahl et al. 2010; Walker et al. 2011).

Research in the field of Computer-Supported Collaborative Learning (CSCL) has consistently revealed that collaboration can foster students' domain knowledge and also domain-general skills, such as argumentation, critical thinking and problem-solving ability (Dillenbourg 1999; Weinberger and Fischer 2006). The key learning mechanism in collaboration is peer interactions; however, there is no guarantee that these interactions will actually occur (Dillenbourg and Tchounikine 2007). For instance, although the 'collaboration contract' implicitly assumes that all partners will contribute to a problem-solving activity, this is not always the case. Research in CSCL has consistently reported on this issue and proposed various methods for increasing the probability that productive peer interactions will actually occur (Dillenbourg 1999). One such method to guide the collaborative activity and foster

peer interactions is the use of a didactic scenario or “collaboration script” (Kobbe et al. 2007).

Following what seems to be another promising path, several research groups have started exploring the use of conversational agents to trigger productive forms of peer dialogue and scaffold students’ learning in a CSCL context. The key objective of these efforts is to design the agent’s behavior in a way that leverages constructive conversational peer interactions (e.g. argumentation, explicit explanation, conflict resolution, mutual regulation) (Fischer et al. 2013). The results have been quite encouraging so far, indicating that the use of conversational agents for providing collaborative learning support can significantly enhance the learning outcomes (e.g. Chaudhuri et al. 2008; Dyke et al. 2012; Kumar et al. 2007b).

Many of the studies above explore the impact that conversational agent interventions may have on the quality of peer dialogue and, consequently, on group and individual learning outcomes. For example, the use of tutorial dialogue agents that encourage peers to engage in directed lines of reasoning, also known as knowledge construction dialogues, have been reported to significantly improve learning outcomes (Chaudhuri et al. 2008; Kumar et al. 2007b). In addition, it was revealed that the unsolicited supportive prompts of an agent can be more effective than solicited supportive prompts (Chaudhuri et al. 2009).

Following a similar rationale, Walker et al. (2011) investigated the role of a conversational agent system that displayed reflective prompts during a reciprocal peer tutoring scenario, where two students took turns tutoring each other. The results showed that the adaptive agent support led students to increase the conceptual content of help in their utterances. Wang et al. (2007) also found that students interacting with a conversational agent during a collaborative brainstorming activity produced substantially more ideas than individuals brainstorming with a human peer. However, one should also keep in mind Kumar et al. (2007b) observation that collaborating peers do not always interact productively with the conversational agents, often ignoring their prompts.

Dyke et al. (2012) and Stahl et al. (2010) explored the use of conversational agents that scaffold collaborative learning discussions through an approach called “Academically Productive Talk” (Michaels et al. 2008). The agents employed a number of supportive moves, which have proven helpful when used in classrooms by teachers. The results provided evidence that one form of support, named “Revoicing”, which encourages students to clarify and/or expand their answers, had a significant positive effect on learning and resulted in higher quality group explanations (Dyke et al. 2012).

Another series of studies focused on the benefits of using conversational agents employing social interaction strategies in collaborative learning situations. Once again, results showed that agents engaging in off-task as well as task-related conversations can enhance learning by efficiently managing students’ attention and focus on the task (Ai et al. 2010a; Gulz et al. 2011; Kumar et al. 2010). Kumar et al. (2007a) also reported a strong positive effect on the attitude that students displayed toward the socially capable agents. In addition, it was found that idea generation productivity can be effectively supported by a conversational agent that makes both instructional and socially-oriented interventions (Kumar et al. 2010).

Nevertheless, along with the abovementioned encouraging outcomes, important research questions have arisen concerning conversational agents’ practical design and

implementation issues in a CSCL context. For instance, which learner should the agent address a question or a comment to when scaffolding a group? How should the agent identify when it is being addressed and which learner is the addressee to the agent (Kumar and Rosé 2011)? What should be the balance of the on-task and off-task conversations between the student and the agent (Gulz et al. 2011)? What types of collaborative problems are suitable for this type of tutoring (Harrer et al. 2006)? What should the teacher's role be when using an agent in a collaborative learning setting? And finally, are the effort and expense required to develop such a system worth the outcome (Walker et al. 2011)?

### Productive Dialogue

A key pattern emerging from the above studies is that an agent's intervention in a collaborative setting can be beneficial for the learners if it can support a sustained 'productive peer dialogue'. However, identifying what a productive peer dialogue is, regardless of the domain, is certainly not a trivial task. In the past, many theorists explored the features that can account for a productive conversation from both a cognitive and socio-cultural perspective (Stahl and Rosé 2011). Despite their differentiation in conceptualization and terminology used, as in "group cognition" (Stahl 2006), "uptake" (Suthers 2006) or "transactivity" (Weinberger and Fischer 2006), their research findings illustrated some consistencies originating from the Vygotskian (Schwartz 1998) and Piagetian (De Lisi and Golbeck 1999) frameworks. The core idea was that knowledge construction during a dialogue occurs through a series of steps, where mental models are articulated, shared, mutually examined and potentially integrated (Sionti et al. 2012; Stahl and Rosé 2011).

The above process lies in the heart of the transactivity theory proposed as a basic theoretical construct for explaining collaborative knowledge construction (Noroozi et al. 2013). Berkowitz and Gibbs (1983) define the term transactivity as "*reasoning operating on the reasoning of the other*". According to Ai et al. (2010b), transactive contributions are arguments constructed upon the previously expressed reasoning of self or others. Likewise, Fischer et al. (2013) indicate that a discourse is considered to be transactive when learners use their partners as their resources, building on their earlier contributions.

In the CSCL domain, the use of transactivity has been repeatedly highlighted as a valuable indicator of the learning taking place in group discussions (Sionti et al. 2012). Teasley (1997) found that transactive contributions are positively correlated to learning outcomes in collaborative problem-solving settings. Similarly, Chi (2009) reported that learning activities involving students in using one another as information resources, building on each other's thoughts, are associated with better learning outcomes as compared to other types of activities. In line with these findings, Stegmann et al. (2011) indicated that a higher level of transactivity can significantly enhance individuals' domain knowledge. Moreover, Noroozi et al. (2013) revealed that transactivity can increase the quality of both the group and the individual problem solutions produced in a problem-solving activity, while both the externalization of students' own knowledge and the elicitation of their partner's knowledge can facilitate collaborative learning.

Another approach consistent with the transactivity perspective is the framework of Academically Productive Talk (APT), which considers social interactions to play a prominent role in inducing beneficial mental processes (Resnick et al. 2010). According to APT, a very effective teaching strategy is to help learners externalize their thinking by “*sharing their reasoning out loud*” (Michaels et al. 2008). Indeed, a major issue is that some learners tend to avoid making their perspectives explicit to the group so that a common ground can be negotiated (Weinberger et al. 2007). This is especially true in written dialogue where the externalized students’ representations can serve as explicit references that facilitate peer interactions and grounding processes (Papadopoulos et al. 2013; Oehl and Pfister 2009). Dillenbourg and Jermann (2007) also underline that the existence of mutual explicit references between learning partners is essential for the success of their collaboration.

### **Mentorchat: A Design Proposal**

Against the above background, we argue that the development of a conversational agent to support learners in collaborative learning, should take into account the current research evidence on:

- The value of agent intervention (prompting) in helping learners sustain a ‘productive peer dialogue’;
- ‘Transactive dialogue’ as a form of productive peer dialogue.

Furthermore, it is our position that the agent should not be designed as teacher replacement, but should serve as an intelligent tool enhancing the impact of the collaboration strategies employed by the teacher, who always remains responsible for the orchestration and scaffolding of the educational activity.

Since it is well documented that peer interaction is a key learning mechanism in collaborative learning, we argue that agent design should not necessarily focus on exhaustively modeling each learner’s understanding by using complex knowledge structures for each different domain. Instead, we suggest that attention should be given on (a) identifying efficient techniques of modeling and triggering constructive peer interactions through appropriate agent interventions (prompts), and (b) promoting the “teacher-as-activity-designer” perspective, by enabling teacher-led configuration of agent behavior (for example, domain modeling and defining agent intervention rules).

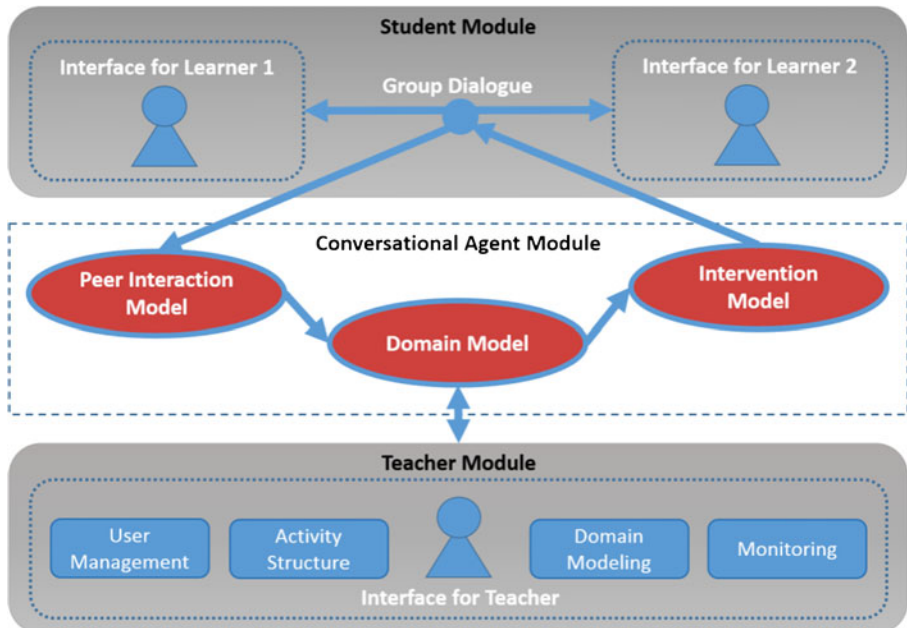
In this way, we expect to develop a conversational agent system (a) of considerable pedagogical value, which will engage peers in a highly transactive dialogue by delivering well-targeted interventions, (b) of a significantly lower development cost, as it will not necessitate hardwiring an explicit domain model, thus being reusable, (c) one enjoying broader acceptance by teachers as being teacher-configurable. We consider that this line of work draws away from the typical design of dialogue-based dynamic support systems, which are tailored for specific learning populations and domains. Indeed, we believe that this approach can lead to design specifications of conversational agent systems broadening their perspective of use as configurable domain-independent intelligent tools with a view to scaffolding diverse student populations, collaborating in various instructional domains.

Following this perspective, we have designed and developed a dialogue-based ITS system called MentorChat. In the next section, we present the MentorChat system architecture, explain our research motivation and provide a list of what we consider as important research questions. In this work, we present a pilot study (as an initial step in this research direction) focusing on what we call “the agent intervention mode”, that is, which learner the agent should address when making an intervention into the peer discourse.

### Mentorchat: System Architecture Overview

MentorChat has been developed as a cloud-based application that requires a small amount of time to be configured and can be used in different learning contexts. The system is available for both the English and the Greek language and uses web technologies, such as MySQL, PHP, AJAX, HTML5 and CSS3. The overall architecture of MentorChat consists of three core modules: the student, the conversational agent and the teacher module (Fig. 1).

Students may login to MentorChat (student module) to participate in synchronous collaborative activities (text-based chatting). An activity can include many online discussion sessions (phases), each on a different topic. In each phase, students in a group are asked to discuss and submit their joint answer to an open-ended domain question. During peer discourse, the MentorChat conversational agent (conversational agent module) is responsible for tracking and analyzing the group dialogues (peer interaction model), identifying opportunities for intervention. Based on the information provided by the teacher (teacher module), which essentially is a domain



**Fig. 1** The abstract system architecture

knowledge representation (domain model) relevant to the topic of discussion, the agent decides when to initiate a supportive intervention. To accomplish this, the agent implements an intervention strategy and presents prompts to the learners in an appropriate manner (intervention model).

## Student Module

Figure 2 displays a screenshot of the MentorChat interface for learners. At the top left corner of the screen, there is a progress bar that indicates both the current phase and the total number of phases of the MentorChat activity (Fig. 2, A). Next to it, a discussion topic is presented to the learners (Fig. 2, B), followed by a chat frame (Fig. 2, C). Also, in the left section of the interface (Fig. 2, D), there is a users' list that displays the online group members participating in this phase. Next to each username, there is an icon representing the role of the user in the current phase of the activity.

When learners login to MentorChat using their web browser, they are forwarded to a private discussion channel according to their assigned group. Subsequently, a conversational agent provides the students with guidance pertaining to their role and task in the current phase of the activity. The interface employs an animated human-like two-dimensional avatar to serve as the agent representation (Fig. 2, E). In addition, a Text-To-Speech engine is utilized to direct the attention of users and enrich the agent modality.

The conversational agent also attempts to efficiently manage students' attention by employing an "Attention Grabbing" strategy, which we have designed based on the work of Kumar and Rosé (2011). According to this technique, when the agent detects an opportunity for support, it does not interrupt students' discussion by immediately making an intervention. It first displays a question mark icon above its avatar and only after some time has passed will the agent's message appear. In that way, the system prepares learners for the intervention to follow, turning their attention to the agent.

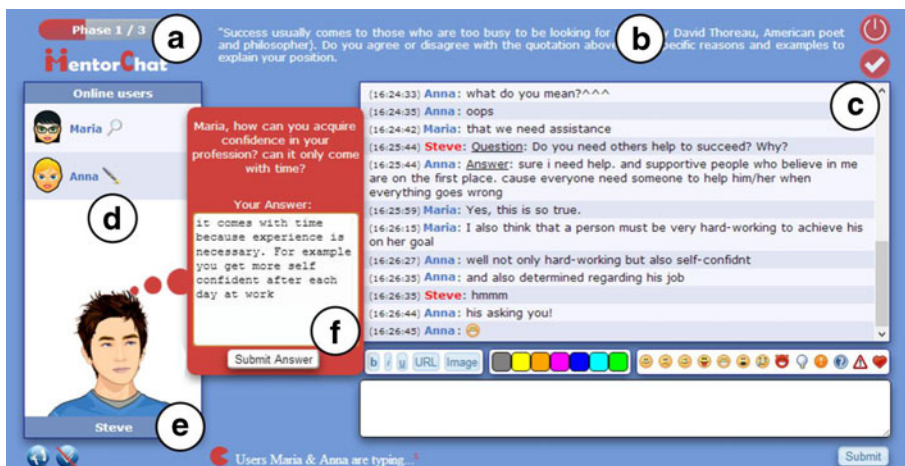


Fig. 2 A screenshot of MentorChat student interface



Furthermore, in contrast to the design of the previous MentorChat versions, in which all agent's messages were presented in the main chat window (Fig. 2, C), the current system version (v.2.5) displays the agent interventions in a bubble that dynamically appears next to the agent avatar (Fig. 2, F). The rationale behind this design decision was derived from the findings of our previous study and responded to shortcomings we identified (Tegos and Demetriadis 2012). More specifically, we observed that students sometimes (a) ignored the agent interventions until they had completed their on-going interaction, and (b) if they did decide to respond, they sometimes had trouble scrolling up the chat list to locate the last agent intervention after some time had passed. As a result, we decided to implement the above agent intervention interface to enable students to have constant access to and respond to the agent's question whenever they felt ready. This mechanism was also found to be helpful in identifying which student of the group has provided a direct answer to the agent. After an answer has been submitted, the bubble fades out and both the agent question and its corresponding answer appear in the chat frame.

### Teacher Module

A teacher can configure MentorChat to design and deploy a collaborative activity through dividing learners in groups, assigning roles to learners, setting up the sequence of the collaborative tasks, and modeling the activity domain, which will be used by the agent to scaffold learners' discussions. In addition, the teacher has the ability to monitor learners' interactions while the activity takes place. All these functions are accessible via the following administration panels: a) the user management, b) the activity structure, c) the domain modeling, and d) the monitoring panels (Fig. 1).

#### *User Management Panel*

The user management panel is responsible for the administration of all participants of the collaborative activity. Using its interface, a teacher can add, delete or edit information such as the username, the password or the group of students. In addition, the teacher can select the roles that will be assigned to each user at the beginning of the activity.

#### *Activity Structure Panel*

The activity structure panel provides the ability to manage the phases of the collaborative activity. In particular, the teacher can create a phase by entering an open-ended domain question or a debate to be resolved (Fig. 2, B). As mentioned before, a phase represents the topic of discussion that encourages students to collaborate in order to reach a common agreement and submit the final answer of their group. The teachers are not restricted to use only text in the discussion topics but they can also insert hyperlinks, images or even videos to present specific concepts to their students in a more comprehensible and engaging manner.

Domain Modeling Panel

The domain modeling panel is essentially an authoring tool that enables the teacher to model the domain and, through that, the agent interventions. It should be noted that learners can only see the final agent intervention appearing on the screen and not the authoring panel. Modeling the domain is possible in two different ways, using either (a) the automatic or (b) the manual modeling method.

The automatic modeling method is based on a concept mapping interface (Fig. 3, A). Through this interface, the teacher can create a concept map by entering a set of ‘facts’. In a basic scenario, a fact is a simple sentence which consists of a subject, an object and a verb. The system interprets these elements as a relationship (as stated by the verb) between two concepts (as stated by the subject and the object), and stores them in a computational format. More specifically, when the teacher enters a fact, the system renders it dynamically, visualizing the nodes and their relationship as a concept graph. This graph constitutes a domain knowledge representation, which is utilized by the agent for implementing its interventions in various ways. For instance, if the system detects that students are talking about two relevant domain concepts, the agent may ask students to comment on whether or how these concepts relate to each other. In this manner, based on the teacher’s graph, which might actually include some incorrect relationships, the agent may ask learners “do you agree or disagree with the following statement: ‘success leads to happiness’? Comment on that”. The automatic modeling method is currently a work in progress that will be discussed and assessed in a future study.

The manual modeling method enables teachers to gain low-level control of the agent’s behavior via a set of rules that determine the agent interventions. Typically, a

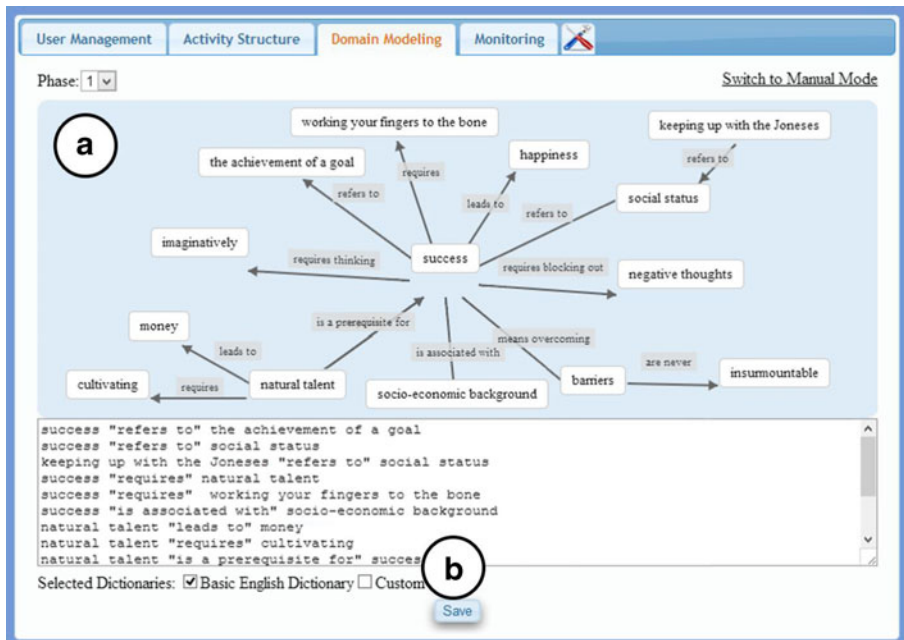


Fig. 3 The interface of the automatic modeling method

teacher-defined rule, which is applicable for a selected phase, can consist of three parts: an ‘event’ (mandatory), a ‘condition’ (optional) and an ‘intervention’ (optional). The event is a linguistic pattern (e.g. a key word) that serves as an opportunity for the display of an agent intervention if the defined condition is met (e.g. after students have discussed about a specific concept). The intervention, which is the only part of the rule displayed to the students, may include not only text but also system operators or other forms of multimedia content such as images, videos or hyperlinks. The operators are system-reserved words that either activate special agent functions, such as facial expressions (e.g. “\_sad\_”, “\_surprised\_”), or serve as variables (e.g. “\_timestamp\_”), which are replaced with the corresponding text when the intervention is delivered.

Moreover, in the manual modeling method, a teacher may decide to omit entering the intervention part of a rule, allowing the system to automatically generate the intervention from scratch. In this case, the agent draws upon the available system dictionaries and the teacher-defined triggering event - for example the presence of a key concept in the students’ dialogue - to dynamically synthesize the agent intervention. The intervention is formed based on either one of the following intervention types: a) encouraging students to explicate their reasoning regarding the key concept (e.g. “Why do you think that *X* is important?”) and b) asking students to apply their own reasoning to their partner’s reasoning (e.g. “Do you agree or disagree with your partner about *X*? Why?”). The selection of these two supportive actions was based on the framework of APT, which refers to them as ‘productive talk moves’ (Resnick et al. 2010).

The authorability of domain models in rule-based dialogue management systems is generally considered to be associated with the sophistication and efficiency of the conversational agents (Sagae et al. 2011). In an attempt to reduce teachers’ workload and support authors of different background (e.g. novices or experienced), we have decided to implement both a higher-level authoring method, utilizing concept mapping as a powerful graphical tool for organizing and representing knowledge (Novak and Cañas 2008), and a rule-based method that enables teachers to manage lower-level details regarding the agent behavior (Sagae et al. 2011). Our system does not enable the identification of multiple dialogue acts in a single student’s utterance as other dialogue management systems do, such as the classifier-based approach (Morbinì and Sagae 2011); still, we consider our approach to offer significant advantages towards the reusability of teacher-defined domain models. This is accomplished by providing access to a library of pre-built concept maps and behavioral rules that can facilitate the domain modeling process.

### *Monitoring Panel*

The monitoring panel allows the teacher to observe the group discussions in real time. There are several filters, such as the student’s name, the date-time or the discussion topic, which enable the teacher to isolate and focus on specific fractions of students’ discourse. In addition, the monitoring panel employs a set of visualized interaction analysis indicators. These are categorized according to their focus on either individual learners (e.g. number of posts, post frequency) or groups (e.g. on-task rate, number of agent interventions). Some interaction analysis indicators have a comparative role

and provide information about the interactions taking place during the activity (e.g. which the most active participants are, which group discussions have covered all the teacher-defined key topics).

In addition to the above administration panels, there is also a general settings panel, which enables the teacher to customize specific options of the educational activity. Particularly, the teacher can alter features, such as the name of the conversational agent, the role instructions, the activity timeframe, the system dictionaries and many others.

## Conversational Agent Module

The architecture of the conversational agent module is based on three models: a) the peer interaction, b) the domain and c) the intervention models (Fig. 1).

### *Peer Interaction Model*

The peer interaction model analyses and stores the most important aspects of students' interactions. More specifically, this model is responsible for identifying and recording all the interactions that could serve as opportunities for a potential intervention. Therefore, by tracking the topics of discussions, the peer interaction model forms the knowledge representations of both the group and the individual learners. These representations are stored in an appropriate format so that they can be easily processed by another computational model.

### *Domain Model*

The conversational agent module exploits the domain model arising from the concept map or the set of rules prescribed by the teacher in the domain modeling panel. In each phase, the teacher-defined domain model (i.e. knowledge representation) is compared with the corresponding individual's or group's knowledge representations to determine whether an agent intervention is appropriate.

To perform the comparison of these knowledge models, MentorChat utilizes a suffix-stripping stemming algorithm in conjunction with a set of appropriately structured dictionaries. The stemmer used is based on two slightly revised PHP versions of the Porter stemming algorithm (Porter 1997) and the Saroukos stemming algorithm (Saroukos 2008) to reduce the inflected or derived words to their stems or root forms. On the basis of this information, MentorChat is then able to use the available dictionaries in order to recall the corresponding synonyms and antonyms. These are essential for (a) the accuracy of the pattern matching algorithms used in MentorChat and (b) the natural language generation required for the synthesis of the agent interventions. The latter system feature is also enhanced by a large pool of stored phrases that can be used to increase the diversity of the agent interventions made.

The system dictionaries have a flexible structure and can be managed through the interface of a dictionary management panel. The creation of a dictionary is a relatively simple process, which can be done by uploading a spreadsheet file (.csv or .xls) with specific headers (e.g. ID, Word, Synonyms and Antonyms). Some dictionaries can be designed for general purpose, and, thus, be used in various

activities, while others can include domain-specific terminology, complementary to that of the general ones, which may not suffice for a particular task. Hence, whenever the teacher models the domain of the activity (via the domain modeling panel), he/she can select which of the listed dictionaries will be used by the system to enhance agent performance (Fig. 3, B). For instance, a teacher organizing an activity about a specific instructional domain can readily create a new dictionary that includes a list of synonym terms relevant to the task. Thus, if the teacher selects the domain-specific dictionary during the setup of the activity domain model, he/she enables MentorChat pattern matching algorithms to detect all the terms that carry the same meaning for the specific activity phase.

### *Intervention Model*

After the agent domain model has indicated that a specific intervention is appropriate, the intervention model is responsible for making the final decision about whether or not an action needs to be taken. This is possible by performing a number of simple checks which ultimately determine if a supportive intervention should be made.

A part of this sequential checking procedure is the estimation of the time passed since the last agent intervention. This timeframe should be over the specified threshold to avoid excessive interference from consecutive agent interventions. Additionally, before intervening in students' discourse, the agent needs to ensure that its previous intervention has already been responded to. Eventually, if all conditions are met, the intervention model must determine the target of the specific intervention. For instance, the agent intervention may address the entire group or just a specific student. In fact, this design choice constitutes an interesting research question to be examined in a following section.

## **Research Motivation and Questions**

Currently, we have set up some broad research directions relating to:

- *Pedagogical effectiveness of the agent*; explore such issues as the level of agent intelligence and modeling required (including, for example, optimization of the peer interaction and the intervention model) so that a transactive dialogue is sustained, resulting in improved learning outcomes.
- *Acceptance by students and teachers*; research usability issues, facilitate the teacher-led domain modeling process and investigate what teachers would consider as beneficial use scenarios for such a system in their courses.

In this work, we specifically present a pilot study in the CALL domain, exploring some issues relating to the pedagogical effectiveness of the agent as well as the students' acceptance and usability of the system. As we have emphasized elsewhere (Magnisalis et al. 2011), an important aspect in the design of collaboration support systems is how the system supportive intervention is presented to the learners. Walker et al. (2009) have proposed that the system intervention may vary, employing either a direct or indirect approach in presenting the prompts. "Direct" means that the system directly addresses the learner who is considered to need support (i.e. the 'weak'

partner), while “indirect” refers to the system addressing the other partner, who is supposed to follow the strategy suggested in order to help the ‘weak’ partner.

Based on the above considerations, the main objective of the study is to explore whether the ‘weak’-directed intervention mode (WDI: the agent addresses one particular student – the ‘weak’ partner) can be more beneficial than the undirected intervention mode (UI: the agent addresses the whole group and not any particular student). We think of this research question as a special case of the broader research issue explained above, that is, the effectiveness of differently presented support intervention techniques in the context of a collaborative activity. We expect the outcome of the study to inform researchers and designers on the pedagogical value of implementing any of these specific agent intervention modes.

## Pilot Study

### Method

#### *Instructional Domain*

The study was conducted in the domain of language learning. Generally, Computer-Assisted Language Learning (CALL) is considered as an ‘umbrella term’ that embraces the use of a wide variety of ICT applications for the acquisition of a foreign language (Levy and Stockwell 2006). The main idea behind most of the language learning applications is the combination of pedagogical theories, such as cognitivism and constructivism, with second language acquisition theories, such as Long’s interaction hypothesis (Long 1996).

Second Language Acquisition (SLA) refers to the process of learning a language other than the mother tongue (DeKeyser 2007). According to Long (1996), SLA is strongly facilitated by using the target language in learners’ interactions. Indeed, the importance of conversational interactions in second language learning has been repeatedly highlighted by researchers (e.g. Mackey 2007).

Former studies have indicated that the use of conversational agents in language learning settings can be quite helpful (Engwall 2008; Fryer and Carpenter 2006; Hjalmarsson et al. 2007; Massaro et al. 2006; Shawar and Atwell 2007; Wik 2011; Wik and Hjalmarsson 2009). Learners who interact with animated conversational agents tend to become more motivated and enjoy their interaction with the system (Fryer and Carpenter 2006; Shawar and Atwell 2007). Indeed, the versatility of conversational agents, which are able to enact different personalities and roles, can increase the interest of students and, thus, the time they spend on the learning task (Wik and Hjalmarsson 2009). Another benefit of embodied conversational agents is their potential to leverage features, such as facial expressions and speech synthesis, allowing them to communicate with language learners using multimodal messages (Hjalmarsson et al. 2007; Massaro et al. 2006; Engwall 2008). According to Burnham and Lau (1999), this can be extremely useful since students listening to a foreign language incorporate visual information to a greater extent than people listening to their native language. Furthermore, Wik (2011) investigated the use of a conversational agent as a teaching companion for language learning. His study revealed that

most of the students perceived their interaction with the embodied agent as a useful learning experience that enriched their vocabulary.

Drawing upon the findings of previous studies, most of them emphasizing on individual language learning, we decided to explore the use of our conversational agent in a collaborative language learning setting. Specifically, the study was carried out in the context of a third-year university course named “Practical Course of the English Language”. This course was offered by the Department of English Philology at the Taurida National University of Ukraine.

### *Participants*

The participants were 30 (26 females and 4 males) undergraduate students enrolled in four different classes of the same teacher. Although students included native speakers of Russian, Tartar, Armenian, Ukrainian, and Hebrew, most of them spoke to each other in Russian. Their age ranged from 18 to 21 ( $M=19.23$ ,  $SD=0.15$ ).

### *Materials*

The materials used in the exploratory study consisted of a pilot installation of the MentorChat tool, a post-task questionnaire and an open-ended interview protocol. Given that most participants were females, we decided to use a male agent (Steve) in MentorChat (Fig. 2, E) considering the research findings that using agents of the opposite gender can increase participants’ performance and motivation (e.g. Karacora et al. 2012).

### *Procedure*

Prior to the date of the study, the teacher proposed three open-ended domain issues, each introducing one of the three phases of the activity. All issues were related to the concept of ‘success in life’ (e.g. “what do you believe are the ‘keys to success’?”), which had been previously discussed in classroom. Additionally, the teacher configured the agent behavior via a set of rules. Each rule consisted of a key word or phrase and a relevant agent intervention. The key words and phrases selected by the teacher also included some misspelled or easily confused words based on the teacher’s experience and awareness of students’ most common errors. The interventions were reflective questions that required students to elaborate their discussion and provide a thoughtful, reasoned response (e.g. “Do you think that social background can influence career success? In what way?”).

First, students attended a 10-min presentation of MentorChat key functionalities, which aimed to familiarize them with the interface options and the types of agent interventions. Subsequently, students were randomly assigned to small groups (12 dyads and 2 triads) and asked to use MentorChat to collaborate and provide their joint answers to the three open-ended domain issues. The activity took place in two university computer labs.

At the beginning of every activity phase, the agent posted some messages that increased students’ awareness (e.g. “Janna has just entered phase 3”) and provided guidance on the roles students had to enact. Participants were assigned to either the

role of author or the reviewer. The author was responsible for initiating the discussion and submitting the final answer of the group to the open-ended question at the end of each phase. The reviewer had to propose improvements on author's contributions by commenting, suggesting corrections, introducing additional information and so on. When students submitted their joint answer, they could proceed to the next phase where their roles were reversed. In the case of the triads, there were always two reviewers and one author.

Students were asked to complete the three tasks of the collaborative activity within the 1-h limit. After finishing the activity, students were requested to complete a post-task questionnaire. Finally, 3 days later, students were invited to participate in a focus group session.

### *Compared Conditions*

During the students' discourse, the agent displayed a teacher-defined intervention whenever an associated linguistic pattern was identified. In our study, we examined and compared the effects of two intervention modes: a) the undirected interventions (UI) that targeted the entire group of students (for example, "Can you give an example of a successful person?") vs. b) the 'weak'-directed interventions (WDI) that targeted only the 'weak' student (for example, "Janna, can you give an example of a successful person?"). The agent intervention model was customized to function differently in each activity phase. More specifically, in the 1st phase, half of the teacher-defined interventions were tailored by the system to be WDI while the other half followed the UI formula. The 2nd phase of the activity included only UI while the 3rd phase incorporated only WDI.

As regards the UI (Table 1, row 2), the agent displayed its question to all group members, allowing any of them to respond. Peers were expected to provide a response in a coordinated way (one of them) by typing in the agent answer box (Fig. 4, A), which remained available until an answer had been submitted.

In the WDI condition (Table 2, row 2), although the agent displayed the question simultaneously to all peers, the answer box was disabled for all students (Fig. 4, B2) except for the one specified in the agent question (Fig. 4, B1). According to our WDI mode, the agent addressed the partner of the student who had activated the agent intervention by mentioning key domain concepts. In the following, we refer to the student who introduced these concepts into the discussion as 'S-student' ('S' for 'strong'), whereas the other student(s) are referred to as 'W-students' ('W' for

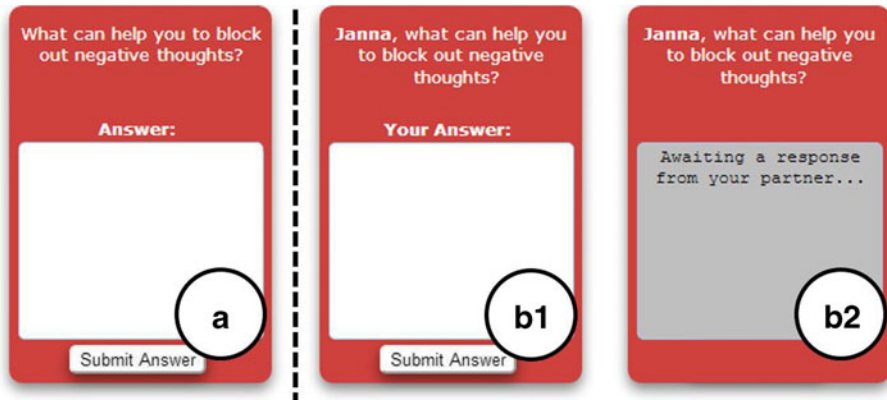
**Table 1** An example of a dialogue containing a UI

---

Alice:	Our life is like a zebra and without bad things we won't understand the importance of the good ones. Nevertheless we should block out the negative thoughts and relieve some of the tension, to take it easy
<b>Agent:</b>	<b>What can help you to block out negative thoughts?</b>
Helen:	Confidence maybe?
Alice:	[Submitted Agent Response] yes confidence is very important because sometimes, when there are external distraction, only our self-belief can help to block all negative thoughts and achieve way beyond we expected

---





**Fig. 4** A: UI interface, B1-B2: WDI interfaces

‘weak’). In Table 2, when Chris (S-student) mentions task-related concepts triggering agent intervention (such as “bad mood” and “think positively”) the agent directs its question to Janna (W-student) to respond. The agent intervention assumption is that, in the current dialogue turn, the W-student (Janna) might have weaker understanding than the S-student (Chris) about the concepts brought up by the S-student. In the case of a three-member group, the agent selects the target of its WDI randomly between the two W-students of the group.

## Measures

### *Post-task Questionnaire*

After finishing the collaborative activity, students completed a student opinion questionnaire, which collected profile data and elicited students’ opinions regarding their collaboration and interaction with the system. The questionnaire included 25 questions: 2 open-ended, 2 multiple-choice, and 21 Likert-scale items. The Likert-scale questions asked students to express their agreement or disagreement on a 5-point scale ranging from 1 (disagree) to 5 (agree).

In order to extract valuable information about the system overall usability and efficiency from the Likert-scale questionnaire items, we came up with the following three indexed variables: ‘perceived usefulness of peer discussions’ (Cronbach’s alpha = 0.74), ‘perceived usefulness of interactions with the agent’ (Cronbach’s alpha = 0.72)

**Table 2** An example of a dialogue containing a WDI

Chris:	When I doubt about my abilities, I’m looking for ways to cheer myself up. I try not to let bad mood become an obstacle in my way to success. It’s important for me to think positively
<b>Agent:</b>	<b>Janna, what can help you to block out negative thoughts?</b>
Janna:	[Submitted Agent Response] I would say having a supportive family and most of all, believing in yourself. you should’ not be afraid of anything because there are no insurmountable obstacles
Chris:	I totally agree with you. No obstacle is ever insurmountable. When times are hard, you keep plugging away towards your goal...

and ‘perceived ease-of-use of system’ (Cronbach’s alpha = 0.71). These metrics were selected considering the principles of Technology Acceptance Model, which is widely used in information systems research to model how users accept and use a technology (Venkatesh and Davis 2000).

Apart from the statistical analysis concerning the above indexed variables, measures of central tendency were also calculated for each variable of the questionnaire, which reflected a high overall reliability (Cronbach’s alpha = 0.88). Furthermore, a series of Pearson product–moment correlation coefficient measures were computed to assess the relationships between the questionnaire variables.

### *Discourse Analysis*

A qualitative analysis of group discussions was performed to draw inferences regarding the effect of the agent interventions. In particular, an adjusted version of the IBIS discussion model was used to identify the types of students’ contributions. The IBIS model was initially designed by researchers (Kunz and Rittel 1970) as a method for structuring discussion activities of collaborative design. It was later proven to be an effective model for analyzing students’ on-line work in collaborative settings (Eggersmann et al. 2003; Liu and Tsai 2008). In the current study, we were specifically interested in measuring the amount of students’ explicit reasoning triggered by agent interventions in the two different conditions (WDI and UI). For this purpose, we adjusted the IBIS scheme to include, among others, the following discussants’ contribution types: ‘Explicit Positions’ and ‘Explicit Arguments’. The overall scheme used in our analysis is as follows:

- *Off-task*: representing contributions irrelevant to the task (e.g. jokes)
- *Repetition*: representing reiterations of previous statements
- *Issue*: representing what needs to be done or resolved as the learners’ task. A contribution of this type could be a question to be answered in order to advance the conversation (e.g. “So, do you think that having someone to support you is important for your success?”)
- *Position*: representing opinions related to the resolution of a raised issue (e.g. “I think having someone to support you is important.”)
- *Explicit Position*: representing positions that explicitly display reasoning about domain concepts (e.g. “Social background can definitely influence career success because some poor families have no opportunities for education.”)
- *Argument*: representing opinions that may support or object to a position (e.g. “I agree with your answer”, “I think you are right.”).
- *Explicit Argument*: representing arguments that (similarly to *explicit positions*) display explicit reasoning about domain concepts (e.g. “I agree! We should mark our mini goals every day because this will help us achieve our final goal”).
- *Management*: representing management-oriented contributions that are useful for the coordination of the activity (e.g. “There is no time left”, “let’s proceed to the next phase.”).
- *Common Understanding*: representing short utterances (usually one- or two-word phrases) used by students to establish a common understanding on the subject of discussion (e.g. “OK?”, “Sure...”, “I see”).

Based on this scheme, student contributions were independently coded by the paper authors. Any disagreement was resolved by discussion.

### *Focus Group*

Students were invited to participate in a focus group session. The focus group protocol was semi-structured and allowed for open-ended discussions. A series of questions were asked to elicit students' opinions about: a) the activity as a whole (likes/dislikes), b) the collaboration with their partners and c) the role of the agent. The qualitative data derived from the discussions were transcribed verbatim and analyzed in search of common themes. An analysis was performed using the open-coding process of the constant comparative method (Corbin and Strauss 1990).

## Results

### *Post-task Questionnaire*

Most students (22 out of 30) were familiar with instant messaging applications. Additionally, students rated their typing skills to be above average ( $n=30$ ,  $M=3.55$ ,  $SD=0.19$ ) using a 5-point scale (1-slow to 5-fast).

Table 3 displays the questionnaire results about the indexed variables related to system usability and user acceptance. A statistical analysis of all questionnaire variables indicated positive student opinions regarding their understanding of the interface functions ( $M=4.77$ ,  $SD=0.21$ ) and of their roles in the activity ( $M=4.77$ ,  $SD=0.11$ ), the system response time ( $M=4.50$ ,  $SD=0.90$ ), the enjoyment derived from their discussions ( $M=4.47$ ,  $SD=0.77$ ), and the perceived quality of the learning experience ( $M=4.30$ ,  $SD=0.99$ ). Also, the majority of the students agreed with statements such as “*the agent questions helped me to recall/find out valuable information about the topics discussed*” ( $M=4.60$ ,  $SD=0.72$ ) and “*I understood the subject better through answering the agent questions*” ( $M=4.47$ ,  $SD=0.57$ ). Slightly lower scores were reported for the performance ( $M=3.60$ ,  $SD=1.13$ ) and helpfulness ( $M=3.86$ ,  $SD=0.94$ ) of the agent's computer-generated voice.

A series of Spearman's rank correlation coefficient measures were also computed to evaluate the relationships between the questionnaire variables. Table 4 reports some of the most significant correlations found.

In an optional question asking students to select their preferred intervention mode, 15 students chose the UI and 9 the WDI mode.

**Table 3** Indexed variable values

Indexed variables	Mean	St. dev.
Perceived usefulness of peer discussions	4.44	0.70
Perceived usefulness of interactions with the agent	4.35	0.61
Perceived ease-of-use of system	4.75	0.27

**Table 4** Significant correlations (\* indicate significance at the 0.01 level)

Correlated variables	Correlation coefficient (r)	Sig. 2-tailed (p)
'frequency of chat usage' - 'discussion interest level'	-0.49	0.02
'role understanding' - 'help received from agent interventions'	0.53*	0.00
'role understanding' - 'improved collaboration skills'	0.39	0.03
'agent's voice helpfulness' - 'enhanced domain knowledge'	0.51*	0.00
'agent's voice helpfulness' - 'help received from agent interventions'	0.65*	0.00
'discussion interest level' - 'enhanced domain knowledge'	0.38	0.04
'discussion interest level' - 'beneficial learning experience'	0.40	0.03
'interface intuitiveness' - 'beneficial learning experience'	0.48*	0.01
'enhanced domain knowledge' - 'beneficial learning experience'	0.77*	0.00

### Discourse Analysis

The discourse analysis revealed a total number of 731 students' contributions, some of which consisted of multiple students' posts. Table 5 depicts the overall results of the coding process conducted for all student groups. Among the other contribution types, it was shown that students produced a total of 190 explicit positions and arguments (Table 5, rows 9 and 10). Such contributions, which display explicit reasoning, play an important role in calculating ERR, a metric described below that measures the efficiency of the intervention model.

Table 6 illustrates the average time groups have spent in each phase of the activity. As expected, the discourse analysis revealed that students spent some extra time in the first phase, providing some off-task contributions that either played a social function or were not directly related to the task. It should also be noted that all groups managed to successfully complete all three phases of the activity.

Among students' discussions, 75 agent interventions were identified, 47 of which were WDI and 28 UI (Table 8, row 2). More specifically, 19 WDI and 8 UI were made during phase 1 ( $n=14$ ,  $M=1.93$ ,  $SD=1.26$ ), 20 UI during phase 2 ( $M=1.43$ ,

**Table 5** The results of the coding process

Contribution type	Total
Off-task	42
Repetition	22
Management	98
Common understanding	29
Issue	96
Position	160
Argument	94
Explicit position	145
Explicit argument	45

**Table 6** The average duration of each activity phase in minutes ( $n=14$  groups)

Activity phase	Mean	St. dev.
1	23.53	3.35
2	18.30	2.78
3	15.23	2.20

SD=0.85), and 28 WDI during phase 3 ( $M=2.00$ ,  $SD=1.30$ ). As for the impact of the two intervention modes, it was shown that 76 out of the 190 explicit positions and arguments found in the students' dialogues (Table 5, rows 9 and 10) were induced by the agent (Table 8, rows 3 and 4). A contribution was identified as agent-induced if it was strongly related to the agent intervention (regardless of the agent intervention mode and the partner who provided the contribution). In this manner, it could be either a direct answer provided via the agent answer box or a following comment relevant to the agent intervention. The coding process of the agent-induced positions and arguments was found to be a relatively easy task since most of the students' messages included explicit references to the agent intervention, such the ones depicted in Table 7 below.

To analyze the impact of the different agent intervention modes on the 'explicitness' of students' contributions, we proceeded to calculate the 'explicit response ratio' (ERR) for both WDI and UI conditions. Measurements were taken from all the three phases of the activity. For each student group ( $n=14$ ), we divided the number of the agent-induced explicit contributions (positions and arguments) by the number of the agent interventions triggered in the group (Table 8). We consider ERR as a statistical measure of the explicit responses generated in the dialogue due to agent intervention. To control for statistical significance, we applied parametric statistics since the normality and the homogeneity of variance criteria were satisfied. A paired sample T-test identified significant difference in favor of the WDI condition ( $n=14$ ,  $M=1.21$ ,  $SD=0.44$ ) as compared to the UI condition ( $n=14$ ,  $M=0.77$ ,  $SD=0.58$ );  $t(13)=-2.70$ ,  $p=0.02$ ,  $d=0.85$ .

Due to the number difference between the two intervention modes (see Table 8, row 2), we also applied a correlation analysis in order to evaluate whether the ERR of the interventions were related to the number of interventions made in each group. The

**Table 7** An example of an agent intervention (WDI) inducing multiple explicit students' responses

<b>Agent:</b>	<b>Sophia, is socio-economic background a factor that can influence career success? In what way?</b>
<b>Sophia:</b>	<i>[Submitted Agent Response]</i> yes I think that it is the factor that can really influence career success. Because families with a poor financial background cannot afford top schools/universities and thus they do not have the same opportunities for high-quality education. Educational background is a crucial in becoming successful
<b>Anna:</b>	No Steve, I disagree with Sophia. According to history, a lot of genius and successful people have achieved more than anyone would have expected by starting at the bottom and working hard. However, I believe that to succeed influential acquaintances are important like in Taras Shevchenko's case...

**Table 8** The impact of the two intervention modes on ‘explicitness’

	WDI mode	UI mode
Number of agent interventions	47	28
Agent-induced explicit positions	42	18
Agent-induced explicit arguments	12	4
Overall Explicit Response Ratio (ERR)	$(42+12)/47=1.15$	$(18+4)/28=0.79$
Average value of Explicit Response Ratios (ERRs)	1.21	0.77

calculated Pearson product–moment correlation coefficients yielded no significant correlations between the quantity and the ERR of the agent interventions.

### Focus Group

Three common themes were identified in the qualitative analysis of the focus group discussions (Table 9). Specifically, students stated that: a) the agent enriched their vocabulary by helping them recall important concepts, previously discussed in classroom (e.g. “*It has been very helpful in remembering the words and phrases we had already discussed during the lectures*”), b) although the computer-generated voice of the agent was helpful, it was not perceived as natural (e.g. “*It is useful and understandable, but can’t be compared to the naturalness of human voice*”), and c) they would prefer to discuss with an English native speaker rather than their classmate (e.g. “*It would be better if I could discuss with a foreign student. A native speaker would be awesome!*”).

## Discussion and Conclusions

In this paper, we have presented the design of MentorChat, a prototype system utilizing a teacher-configurable domain-independent conversational agent to trigger students’ productive dialogue. Additionally, we conducted a pilot evaluation study in the domain of second language acquisition, with the results indicating that the agent WDI mode is more effective in increasing explicit contributions in peer dialogue, as compared to the UI mode. We consider this to be a useful indication regarding the design of conversational agents in the CSCL context.

**Table 9** Qualitative themes

No.	Theme
1	Agent interventions enhanced vocabulary skills
2	The voice of the agent was ‘unnatural’
3	A discussion with a foreign student (especially a native speaker) would be preferable

## Discussion

### *Post-task Questionnaire*

Regarding user acceptance and usability of the system some positive students' opinions were identified. In particular, students perceived the system as an easy-to-use tool and considered their interactions both with their partner(s) and the agent to be helpful for language learning (Table 4). Students enjoyed the discussion with their partner(s) and considered the collaborative activity to be a beneficial learning experience. Moreover, they stated that the agent increased their interest in the discussion and helped them recall important domain concepts, previously discussed in classroom. We believe that the agent questions helped students externalize parts of their domain understanding and practise their language skills by making their thoughts, reasons and ideas explicit. Thus, our study provides additional research evidence regarding the already recorded potential of reflective questions to maintain students' interest, encourage students to think on the content of the lesson, and elicit particular structures of vocabulary items in language learning settings (Richards and Lockhart 1994).

Some significant correlations were also found with regard to the students' answers in the post-task questionnaire (Table 4). Firstly, it was recorded that students who did not frequently use chat applications enjoyed their discussion in MentorChat (Table 4, row 2) more. This is not surprising since students who are unfamiliar with instant messaging applications are expected to be more excited by the use of such a communication tool.

Furthermore, students who better understood their roles during the activity (author/reviewer) felt that the agent interventions helped them (Table 4, row 3) and improved their collaboration to a greater extent (Table 4, row 4). We interpret this as additional research evidence supporting the view that the effectiveness of role playing in CSCL is reasonably linked to the success of the entire collaboration process (Kobbe et al. 2007).

In addition, participants who claimed to have been helped the most by the agent's voice also considered that (a) the agent interventions were more helpful (Table 4, row 5) and (b) the activity enhanced their domain knowledge (Table 4, row 6). It seems that, despite the major challenges that Text-To-Speech (TTS) software face (Taylor 2009), the utilization of speech synthesis techniques constitutes a crucial design factor, even for text-based dialogue systems.

Another correlation indicated that students who were more interested in their peer discussion, perceived the collaborative activity as a more beneficial learning experience (Table 4, row 5), which improved their domain knowledge (Table 4, row 6). Finally, it seems that the interface intuitiveness is positively correlated to the extent to which learners considered the activity as beneficial for learning (Table 4, row 7).

### *Discourse Analysis*

The aim of this pilot study was to compare the effectiveness of two types of agent interventions: a) undirected intervention (UI), that is, the agent addressing all group members, and b) 'weak'-directed intervention (WDI), i.e., the agent questions

targeting only the ‘weak’ or W-student in each case. As explained, in WDI the S-student is always the student who triggers the agent intervention by mentioning key domain concepts (according to rules set by the teacher), and the agent directs its intervention (prompting question) to the other student (W-student).

In order to assess the impact of agent intervention mode on students’ dialogue, we proceeded to a discourse analysis, modeling each agent intervention as a generator of  $N$  explicit responses and introducing a new metric called ERR (explicit response ratio). The ERR is basically a statistical measure of how many explicit responses (positions and arguments) are triggered by an agent intervention. In this manner, ERR can range from zero - if no explicit responses are stimulated - to more than one if multiple explicit responses are stimulated by the agent intervention. The analysis showed that the WDI condition triggered considerably more responses displaying explicit reasoning, as compared to the UI condition (Table 8). Attention is called to the fact that this study only measures the ERR *hypothesizing* that high ERRs are connected to highly transactive forms of peer dialogue and to a high level of learning outcomes. However, the latter are assumptions that need to be further investigated.

A possible explanation for this higher ERR effect of the WDI mode may be the fact that WDI applies some beneficial situational constraints affecting the configuration of peers’ internal scripts and, consequently, peers’ behavior. The term “internal script” refers to the knowledge an individual has about a recurring collaborative situation and, more specifically, to the internalization of collaborative practices as collaboration skills and general cognitive strategies. According to Fischer et al. (2013), in a collaborative activity learners’ understanding and actions are guided by internal scripts, which can be dynamically configured multiple times during the activity. Kollar et al. (2007) indicate that the development of well-configured internal scripts can improve subject matter knowledge acquisition and increase the quality of the arguments produced in collaborative settings.

In our study, we argue that the WDI mode has facilitated the configuration process of students’ internal scripts by inducing situational constraints that reduced the probability of dysfunctional configuration of internal script components. As we are in a position to claim (after carefully extending our peers’ discourse analysis to identify patterns of students’ behavior after an agent intervention occurred), there was little to no coordination between partners in the case of the UI mode. Most of the times, students did not talk to each other about who should submit the answer to the agent. In fact, some students rushed to respond first. On the other hand, in the WDI condition, we identified a beneficial interaction pattern. In 7 (out of 14) groups, we observed that when an agent intervention targeted the W-student, after the W-student had responded, the S-student also provided an explicitly reasoned contribution as a follow-up response. Although further research with a larger sample is required to extend our understanding about this interaction pattern, we found that the S-student often made explicit his/her opinion too, elaborating on concepts addressed by the agent question. Bearing in mind that this kind of interaction processes can increase the level of transactivity (Ai et al. 2010b), we argue that the WDI mode can promote students’ productive talk more effectively than the UI mode. We expect that the WDI mode can encourage more transactive conversational behaviors as compared to the UI mode, by asking the W-student to provide a reasoned response to questions that originated from previous contributions of the S-student. Similar views on how to



augment the transactive quality of peer dialogue, by directing one student to reason based on the other's contributions, have been proposed by other researchers as well (Sionti et al. 2012). Worth mentioning is the fact that most students selected the UI as their preferred intervention mode in the post-task questionnaire, which can be probably explained by the higher level of freedom experienced by students in the UI condition.

In general, we believe that the effectiveness of agent interventions can be improved by choosing a more appropriate level of coercion and inducing students to follow a specific protocol in their interaction with the agent. Therefore, enabling the system to regulate such student-agent interaction processes by itself may be more efficient than just encouraging and expecting the students to play that role themselves.

### *Focus Group*

During the focus group session, almost all students stated that the agent interventions enhanced their vocabulary skills through the question-answering process (Table 9, row 1). When they were asked why, some students mentioned that the agent questions introduced a new discussion perspective that made them expand their thinking and use new words.

Another theme identified was associated to the agent computer-generated voice (Table 9, row 2). Although the students considered the agent voice as helpful for the task, they commented that it was 'unnatural' and 'robot-like', mentioning that "*it could certainly be improved*". This result is in line with Veletsianos (2012) findings, who also reported a similar learners' dissatisfaction with the computer-generated voice of the agent. Despite the above, it should be noted that, in our study, the agent voice was not perceived as a dialogue disruptive feature by the students. On the contrary, some students stated that the agent speech helped them turn their attention to the agent intervention.

Finally, it was recorded that most Ukrainian students of English Philology would prefer to discuss with a foreign student (English native speaker) instead of their classmate (Table 9, row 3). This suggests that in the CALL domain a more engaging (and possibly of higher learning impact) conversational agent, should be 'disguised' as and possess the conversational skills of a target-language native speaker. However, this type of agent design analysis is beyond the scope of this work.

### *Study limitations*

This study has certain limitations such as the short length of the activity and the small number of participants. Given that this is a pilot study, we consider it to provide valuable insights to be verified by conducting a more robustly designed research study in the future. Another limitation might be due to the ordering effects emerging from the specific phase sequencing. More specifically, placing the WDI intervention mode in the third (and final) phase of the activity may have caused a bias in favor of the WDI mode (that is, students being familiar with the agent intervention might have been more active in producing explicit responses during the third WDI-based phase). Although such an ordering effect cannot be excluded, we do not believe that ordering

effects have a major impact on the outcomes, considering that: first, measuring the impact of the two modalities was also based on data recorded during the first activity phase, indeed 19 (out of 47) WDI (40 % from all WDI) already appeared in the first activity phase, when there was hardly any particular ordering effect; second, a demonstration of the two intervention modalities took place before starting the activity, therefore the students were, to a certain extent, familiar with both modalities right from the beginning.

## Conclusion

Despite the limitations discussed above, we consider the results of this pilot study as encouraging and promising. Our findings indicate that the utilization of such a teacher-configurable conversational agent can be both feasible and beneficial for language learning. In addition, the study provides some preliminary evidence in favor of, what we call, a ‘weak’-directed agent intervention mode (with the agent addressing only a specific group member). This modality seems to engage students in a more productive dialogue (as measured by the ERR variable) than an undirected intervention mode. Undoubtedly, more research is needed to better understand how students’ productive talk can be facilitated by efficiently modeling the student-agent interactions in collaborative situations. We consider that MentorChat creates some interesting opportunities for targeted research in the area of conversational agents for collaborative learning.

**Acknowledgments** The authors would like to thank Prof. H. V. Polkhovskaya, Head of the Department of English Philology at the TNU, for her critical role in organizing the pilot study presented in this paper.

## References

- Ai, H., Kumar, R., Nguyen, D., Nagasunder, A., & Rosé, C. P. (2010a). Exploring the effectiveness of social capabilities and goal alignment in computer supported collaborative learning. In *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 134–143). Berlin: Springer.
- Ai, H., Sionti, M., Wang, Y. C., & Rosé, C. P. (2010b). Finding transactive contributions in whole group classroom discussions. *Proceedings of the 9th International Conference of the Learning Sciences* (pp. 976–983). International Society of the Learning Sciences.
- Aleven, V., Ogan, A., Popescu, O., Torrey, C., & Koedinger, K. (2004). Evaluating the effectiveness of a tutorial dialogue system for self-explanation. In *Proceedings of the 11th International Conference on Intelligent Tutoring Systems* (pp. 29–33). Berlin: Springer.
- Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3559–3565.
- Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly*, 29, 399–410.
- Burnham, D., & Lau, S. (1999). The integration of auditory and visual speech information with foreign speakers: The role of expectancy. *Proceedings of AVSP* (pp. 80–85).
- Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. F. (Eds.) (2000). *Embodied conversational agents*. Cambridge: MIT Press.
- Chaudhuri, S., Kumar, R., Joshi, M., Terrell, E., Higgs, F., Aleven, V., et al. (2008). It’s not easy being green: Supporting collaborative “green design” learning. In *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 807–809). Berlin: Springer.

- Chaudhuri, S., Kumar, R., Howley, I., & Rosé, C. P. (2009). Engaging collaborative learners with helping agents. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 365–372). Amsterdam: Ios Press.
- Chi, M. T. (2009). Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105.
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: procedures, canons, and evaluative criteria. *Qualitative sociology*, *13*(1), 3–21.
- De Lisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3–37). Mahwah: Erlbaum.
- Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies*, *52*(1), 1–22.
- DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1–19). Oxford: Elsevier.
- Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning*, *23*(1), 1–13.
- Dillenbourg, P., & Jermann, P. (2007). Designing integrative scripts. In F. Fischer, H. Mandl, J. Haake, & I. Kollar (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 275–301). New York: Springer.
- Dyke, G., Adamson, D., Howley, I., & Rosé, C. P. (2012). *Enhancing scientific reasoning and explanation skills with conversational agents*. Manuscript submitted for publication.
- Eggersmann, M., Gonnet, S., Henning, G. P., Krobb, C., Leone, H. P., & Marquardt, W. (2003). Modeling and understanding different types of process design activities. *Latin American applied research*, *33*(2), 167–175.
- Engwall, O. (2008). Can audio-visual instructions help learners improve their articulation?—An ultrasound study of short term changes. *Proceedings of Interspeech 2008* (pp. 2631–2634).
- Evens, M., & Michael, J. (2006). *One-on-one tutoring by humans and computers*. Mahwah: Lawrence Erlbaum Associates.
- Feng, D., Shaw, E., Kim, J., & Hovy, E. (2006). An intelligent discussion-bot for answering student queries in threaded discussions. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 171–177). New York: ACM.
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, *48*(1), 56–66.
- Fryer, L., & Carpenter, R. (2006). Emerging technologies: bots as language learning tools. *Language, Learning and Technology*, *10*(3), 8–14.
- Graesser, A. C., Chipman, P., Haynes, B., & Olney, A. M. (2005a). AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, *48*(4), 612–618.
- Graesser, A. C., Person, N., Lu, Z., Jeon, M. G., & McDaniel, B. (2005b). Learning while holding a conversation with a computer. In L. PytlíkZillig, M. Bodvarsson, & R. Bruning (Eds.), *Technology-based education: Bringing researchers and practitioners together* (pp. 143–167). Greenwich: Information Age Publishing.
- Graesser, A. C., Jeon, M., & Duffy, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, *45*(4–5), 298–322.
- Grigoriadou, M., Tsaganou, G., & Cavoura, T. (2005). Historical text comprehension reflective tutorial dialogue system. *Educational Technology & Society*, *8*(4), 31–46.
- Gulz, A., Haake, M., Silvervarg, A., Sjödn, B., & Veletsianos, G. (2011). Building a social conversational pedagogical agent-design challenges and methodological approaches. In D. Perez-Marin, & I. Pascual-Nieto (Eds.), *Conversational agents and natural language interaction: Techniques and effective practices* (pp. 128–155). Hershey: IGI Global.
- Haake, M., & Gulz, A. (2009). A Look at the roles of look & roles in embodied pedagogical agents—a user preference perspective. *International Journal of Artificial Intelligence in Education*, *19*(1), 39–71.
- Harrer, A., McLaren, B. M., Walker, E., Bollen, L., & Sewall, J. (2006). Creating cognitive tutors for collaborative learning: steps toward realization. *User Modeling and User-Adapted Interaction*, *16*(3), 175–209.
- Heffernan, N. T., & Croteau, E. A. (2004). Web-based evaluations showing differential learning for tutorial strategies employed by the Ms. Lindquist tutor. In *Proceedings of the 7th International Conference on Intelligent Tutoring Systems* (pp. 491–500). Berlin: Springer.

- Hjalmarsson, A., Wik, P., & Brusik, J. (2007). Dealing with DEAL: a dialogue system for conversation training. *Proceedings of SigDial 2007* (pp. 132–135).
- Jia, J. (2004). The study of the application of a web-based chatbot system on the teaching of foreign languages. *Proceedings of the 15th Conference of the Society for Information Technology and Teacher Education* (pp. 1201–1207). Atlanta: AACE Press.
- Karacora, B., Dehghani, M., Kraemer-Mertens, N., & Gratch, J. (2012). The influence of virtual agents' gender and rapport on enhancing math performance. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 563–568).
- Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A., & Buchan, I. (2012). Active assistance technology for health-related behavior change: An interdisciplinary review. *Journal of Medical Internet Research*, *14*(3). doi:10.2196/jmir.1893
- Kerly, A., Ellis, R., & Bull, S. (2008). CALMsystem: a conversational agent for learner modelling. *Knowledge-Based Systems*, *21*(3), 238–246.
- Kerly, A., Ellis, R., & Bull, S. (2009). Conversational agents in E-Learning. *Proceedings of AI 2008* (pp. 169–182).
- Kinshuk, A. P. (2002). Does intelligent tutoring have future! *Proceedings of the International Conference on Computers in Education 2002*.
- Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., et al. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*(2), 211–224.
- Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction*, *17*(6), 708–721.
- Kumar, R., & Rosé, C. (2011). Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies*, *4*(1), 21–34.
- Kumar, R., Gweon, G., Joshi, M., Cui, Y., & Rosé, C. P. (2007a). Supporting students working together on math with social dialogue. *Proceedings of the SLATE-2007 Workshop* (pp. 96–99).
- Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., & Robinson, A. (2007b). Tutorial dialogue as adaptive collaborative learning support. In R. Luckin, K. R. Koedinger, & J. Greer (Eds.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 383–390). Amsterdam: Ios Press.
- Kumar, R., Ai, H., Beuth, J., & Rosé, C. (2010). Socially capable conversational tutors can be effective in collaborative learning situations. In *Proceedings of the 10th International Conference on Intelligence Tutoring Systems* (pp. 156–164). Berlin: Springer.
- Kunz, W., & Rittel, H. (1970). *Issue as elements of information systems. Working paper 131*. Berkeley: Institute of Urban and Regional Development, University of California.
- Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer assisted language learning*. Mahwah: Lawrence Erlbaum Associates.
- Litman, D. J., & Silliman, S. (2004). ITSPOKE: An intelligent tutoring spoken dialogue system. *Proceedings of the 4th meeting of HLT/NAACL* (pp. 52–54).
- Liu, C. C., & Tsai, C. C. (2008). An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. *Computers & Education*, *50*(3), 627–639.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic.
- Mackey, A. (Ed.). (2007). *Conversational interaction in second language acquisition: A series of empirical studies*. Oxford: Oxford University Press.
- Magnisalis, I., Demetriadis, S., & Karakostas, A. (2011). Adaptive and intelligent systems for collaborative learning support: a review of the field. *IEEE Transactions on Learning Technologies*, *4*(1), 5–20.
- Massaro, D. W., Liu, Y., Chen, T. H., & Perfetti, C. (2006). A multilingual embodied conversational agent for tutoring speech and language learning. *Proceedings of the Interspeech 2006* (pp. 825–828).
- Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297.
- Morbini, F., & Sagae, K. (2011). Joint identification and segmentation of domain-specific dialogue acts for conversational dialogue systems. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Short Papers-Volume 2* (pp. 95–100).
- Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers and Education*, *61*(2), 59–76.

- Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. Technical Report IHMC CmapTools 2006–01 Rev 01–2008.
- Oehl, M., & Pfister, H. R. (2009). Explicit references in chat-based CSCL: Do they facilitate global text processing? Evidence from eye movement analyses. In A. Dimitracopoulou, C. O'Malley, D. Suthers & P. Reimann (Eds.), *Computer supported collaborative learning practices - CSCL2009 community events* (pp. 138–140). Rhodes: International Society of the Learning Sciences.
- Papadopoulos, P. M., Demetriadis, S. N., & Weinberger, A. (2013). "Make it explicit!": improving collaboration through increase of script coercion. *Journal of Computer Assisted Learning*, 29(4), 383–398.
- Porter, M. (1997). An algorithm for suffix stripping. In K. S. Jones & P. Willett (Eds.), *Readings in information retrieval* (pp. 313–316). San Francisco: Morgan Kaufmann Pub.
- Resnick, L. B., Michaels, S., & O'Connor, C. (2010). How (well structured) talk builds the mind. In R. Sternberg & D. Preiss (Eds.), *From genes to context: New discoveries about learning from educational research and their applications* (pp. 163–194). New York: Springer.
- Richards, J. C., & Lockhart, C. (1994). *Reflective teaching in second language classrooms*. New York: Cambridge University Press.
- Rosé, C. P., Torrey, C., & Alevin, V. (2004). Guided exploratory learning versus directed learning in a simulation environment for thermodynamics: A pilot study. *Proceedings of the Workshop on Dialog-Based Intelligent Tutoring Systems held at the ITS 2004* (pp. 43–48). Brazil: Maceió.
- Ruttkay, Z., & Pelachaud, C. (2004). *From brows to trust: Evaluating embodied conversational agents*. Dordrecht: Kluwer Academic.
- Sagae, A., Johnson, W. L., & Valente, A. (2011). Conversational agents in language and culture training. In D. Perez-Marin, & I. Pascual-Nieto (Eds.), *Conversational agents and natural language interaction: Techniques and effective practices* (pp. 358–377). Hershey: IGI Global.
- Saroukos, S. (2008). Enhancing a Greek language stemmer, MSc Thesis, University of Tampere.
- Schwartz, D. L. (1998). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 197–218). Oxford: Pergamon Press.
- Shaw, B. A., & Atwell, E. (2007). Fostering language learner autonomy via adaptive conversation tutors. *Proceedings of Corpus Linguistic Conference 2007*.
- Sionti, M., Ai, H., Rosé, C. P., & Resnick, L. (2012). A framework for analyzing development of argumentation through classroom discussions. In N. Pinkwart & B. McClaren (Eds.), *Educational technologies for teaching argumentation skills*. (pp. 28–55). Oak Park: Bentham Science.
- Sleeman, D. H. & Brown, J. S. (Eds.) (1982). *Intelligent tutoring systems*. New York: Academic Press.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge: MIT Press.
- Stahl, G., & Rosé, C. P. (2011). Group cognition in online teams. In E. Salas & S. M. Fiore (Eds.), *Theories of team cognition: Cross-disciplinary perspectives*. New York: Routledge/Taylor & Francis.
- Stahl, G., Rosé, C. P., O'Hara, K., & Powell, A. B. (2010). Supporting group math cognition in virtual math teams with software conversational agents. In *First North American GeoGebra Conference*. Ithaca, NY.
- Stegmann, K., Weinberger, A., & Fischer, F. (2011). Aktives Lernen durch Argumentieren: Evidenz für das Modell der Argumentativen Wissenskonstruktion in Online-Diskussionen. [Active learning by argumentation: Evidence for the model of argumentative knowledge construction in online discussions.]. *Unterrichtswissenschaft*, 39(3), 231–244.
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: a research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 315–337.
- Taylor, P. (2009). The text-to-speech problem. In *Text-to-speech synthesis*. Cambridge: University Press.
- Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (pp. 361–384). Berlin: Springer.
- Tegos, S., & Demetriadis, S. (2012). Designing a teacher customizable conversational agent to trigger online students' interaction. *Proceedings of the Workshop on Intelligent Support for Learning in Groups held at the ITS 2012* (pp. 11–18).
- VanLehn, K., Graesser, A., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). Natural language tutoring: A comparison of human tutors, computer tutors, and text. *Cognitive Science*, 31(1), 3–52.
- Veletsianos, G. (2012). How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences. *Computers in Human Behavior*, 28(1), 275–283. Elsevier Ltd.

- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science*, *46*(2), 186–204.
- Walker, E., Rummel, N., & Koedinger, K. R. (2009). Beyond explicit feedback: New directions in adaptive collaborative learning support. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds.), *Computer supported collaborative learning practices - CSCL2009 conference proceedings* (pp. 552–556). Rhodes: International Society of the Learning Sciences.
- Walker, E., Rummel, N., & Koedinger, K. R. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer-Supported Collaborative Learning*, *6*(2), 279–306.
- Wang, H. C., Rosé, C. P., Cui, Y., Chang, C. Y., Huang, C. C., & Li, T. Y. (2007). Thinking hard together: The long and short of collaborative idea generation in scientific inquiry. In C. Chinn, G. Erkens & S. Puntambekar (Eds.), *The computer supported collaborative learning (CSCL) conference 2007, Vol 8, Part 2* (pp. 754–763). New Jersey: International Society of the Learning Sciences.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95.
- Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: concepts and assessment. *Learning and Instruction*, *17*(4), 416–426.
- Wenger, E. (2004). Artificial intelligence and tutoring systems. *International Journal of Artificial Intelligence in Education*, *14*, 39–65.
- Wik, P. (2011). The virtual language teacher: Models and applications for language learning using embodied conversational agents (Doctoral dissertation).
- Wik, P., & Hjalmarsson, A. (2009). Embodied conversational agents in computer assisted language learning. *Speech communication*, *51*(10), 1024–1037.