

# Chatbots for Action-Oriented Language Learning Using Elbot to Enhance Conflict-Solving Skills in Learners of German as a Foreign Language

Francesca Mazzilli

Università degli Studi di Bari «Aldo Moro», Italia

**Abstract** The aim of this study is to investigate (i) how learners perceive the non-educational chatbot Elbot as a language learning tool in a formal context (university German course), and (ii) to what extent a set of task-oriented interactions with Elbot influences the learners' behaviour (e.g. lexical choices) during a separate activity. For these purposes, this study outlines a lesson/experiment focused on the enhancement of conflict-solving skills in German as a foreign language. A mixed-methods approach was applied to evaluate Elbot's contribution to the lesson as well as its impact on the learners' choice of basic strategies to solve a conflict. The study concludes with the discussion of the advantages of developing an edu-bot based on an existing non-educational chatbot. In this regard, the research also provides some concrete proposals based on the results of the data analysis.

**Keywords** Chatbots. Conflict-solving skills. Action-oriented language education. Dialogue-based CALL. Elbot. German as a Foreign Language. Academic language education. Edu-bots. Non-educational bots. Chat-based human-computer interaction.

**Summary** 1 Introduction. – 2 Definition and Main Characteristics of Chatbots. – 3 Related Research: Chatbot as Language Learning Tools. – 4 Research Design. – 5 Method. – 5.1 Participants. – 5.2 Data Elicitation and Analysis: Focus on Learners' Perception. – 5.3 Data Collection and Analysis: Focus on Elbot's Influence on Human Behaviour. – 6 Selected Results. – 6.1 Analysis and Discussion: Focus on Learners' Perception. – 6.2 Analysis and Discussion: Focus on Elbot's Influence on Human Behaviour. – 7 Conclusion.



**Edizioni**  
Ca' Foscari

## Peer review

Submitted 2020-05-25  
Accepted 2021-03-23  
Published 2021-03-31

## Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



**Citation** Mazzilli, F. (2021). "Chatbots for Action-Oriented Language Learning. Using Elbot to Enhance Conflict-Solving Skills in Learners of German as a Foreign Language". *EL.LE*, 10(1), 95-116.

## 1 Introduction

The aim of this research is to investigate the impact and potential of non-educational chatbots for the enhancement of conflict-solving skills in the foreign language (L2) in an academic context. In particular, I will focus on following research objectives: (i) identifying how the learners perceive the chatbot Elbot<sup>1</sup> as a language learning tool, and (ii) defining to what extent the interaction with Elbot influences the learners' development of basic strategies to deal with conflicts. In so doing, I will also (i) clarify the position of the chatbot in the didactic triangle, i.e. in the field of mutual relations involving learners, teachers, and foreign language (Kansanen, Meri 1999, 111-12), and (ii) suggest how to concretely develop an edu-bot based on a non-educational chatbot from the point of view of teaching methodology and action-oriented foreign language education.

As "intelligent conversational agents with complex, goal-driven behaviour" (Kerly, Hall, Bull 2006, 178), chatbots find an application in several areas, including e-commerce, entertainment, and most recently education (Bii 2013, 218-20; Istrate 2018, 471-3; Wang, Petrina 2013, 127-8). In particular, the use of chatbots as language learning tools has gained more and more attention since the release of some popular educational chatbots, e.g. Duolingo Bots, which offered a limited set of conversations in German, French, and Spanish on iOS systems from 2016 to 2018 (Mazzilli 2019, 145-7).

In this study, I will focus on the German-speaking version of Elbot, a text-based non-educational chatbot, and on its use in a university German course in order to enhance the learners' conflict-solving skills in the L2, which is one of the greatest challenges of foreign language education both in academic and professional contexts. In fact, dealing with conflicts in a foreign language requires not only a punctual comprehension of the interaction structure, but also the experimentation of several strategies and expressions in order to reach ideally a win-win situation (Heringer 2017, 232-41). This process is much more effective if the experimentation (a proper micro-training) takes place in a stress-free learning environment with non-judging interlocutors. Unfortunately, it is difficult to simulate a conflict under such conditions with human partners, whereas chatbots (even non-educational ones) could contribute to offering an adequate experience for this specific learning goal.

In order to investigate the above-mentioned research questions, I proceeded to (i) clarify the definition and characteristics of modern chatbots, (ii) explore the scientific literature on the use of chatbots

---

**1** Elbot was designed by Fred Roberts in 2006 with Artificial Solutions' technology for entertainment purposes. It is available here: <http://www.elbot.de/>.

for language education (e.g. Bibauw, François, Desmet 2015; Huang et al. 2017; Wang, Petrina 2013) and (iii) design and conduct the lesson/experiment *Konflikte im Alltag* ('conflicts in everyday life') to examine the learners' perception of Elbot and its impact on the learners' behaviour through the collection and mixed analysis of empirical data (learners' responses to a questionnaire about the perception of Elbot as learning tool and selected learners' texts and notes produced during the lesson while and after interacting with Elbot). The results of this study will show (i) in what direction the development of language learning tools based on human-computer interaction should go, and especially (ii) how an edu-bot based on a non-educational chatbot (such as Elbot) and aimed to contribute to a stress-free learning environment should look like.

## 2 Definition and Main Characteristics of Chatbots

As mentioned above, chatbots are "intelligent conversational agents with complex, goal-driven behaviour" (Kerly, Hall, Bull 2006, 178). They are one of the several manifestations of artificial intelligence (AI), which is based on the automation of cognitive processes and on the simulation of human behaviour (Russell, Norvig 2016, 1020). Since we still do not have a strong AI at our disposal, i.e. a thinking system with an individual consciousness, this complex and intelligent behaviour can only be simulated at the moment, still with several critical points on the structural and thematic side, as the following example involving the multilingual chatbot Cleverbot<sup>2</sup> shows:

- (1) NutzerIn: Do you have friends?
- (2) Cleverbot: Not really.
- (3) NutzerIn: Why not?
- (4) Cleverbot: Because I don't like math that much. (Mazzilli 2018, 18)

In this case, even if there are no critical points on the cohesive side (e.g. the correlation 'why'/'because' is correctly applied in this short exchange), the text is not consistent on a thematic level, since Cleverbot (line 4) does not answer the user's question (line 3) in a plausible way. Its response does not correspond at all to the general topic of the conversation (i.e. friendship). In spite of the occurrence of such phenomena in the interaction with chatbots (e.g. *non sequitur*), this has not prevented them from spreading out during the last decades, especially on social media (Lotze 2016, 39-40; Mazzilli 2018,

---

<sup>2</sup> <https://www.cleverbot.com/>.

13). Kerly, Hall, and Bull (2006, 178) identify following characteristics of chatbots, included in their above-mentioned definition:

- *intelligence*: chatbots aim to imitate intelligent communicative behaviour designed for pursuing a specific goal. Since it is de facto still impossible to formulate a consistent definition of intelligence, it is meant here as the capability of (inter)acting in a human-like way (Storp 2002, 1).
- *conversational agent*: as “chatting robot[s]” (Lotze 2016, 39), these systems are chat-based programs which are capable of (i) elaborating texts written by human and non-human users in a natural language, and (ii) reacting to these texts in a natural language (Tewes 2005, 243; Lotze 2016, 39; Storp 2002, 1). Besides, the conversational agents examined in this research should not be mistaken for bad bots and social bots, which can be “directed to attack users (targets) to pursue a variety of latent goals, such as to spread information or to influence users” (Wagner et al. 2012, 41).
- *complexity*: regardless of whether chatbots are available on a web site or as an application that can be installed on an electronic device, their architecture tends to include (Tewes 2005, 248-51): (i) a database consisting of words, clauses, sequences or entire dialogues organised according to specific criteria, (ii) a programming language, and (iii) other components for managing the interaction (dialogue manager etc.) or producing multimodal outputs (graphic interface etc.).
- *goal-drivenness*: chatbots are always developed to pursue a specific goal (informing, entertaining etc.). This goal can be more or less binding in regard to the ‘behaviour’ of the chatbot itself, influencing the variety and the characteristics of its contribution to the interaction (Tewes 2005, 253).
- *behaviour*: some of the most important issues of chatbots (e.g. lack of coherence and credibility in the interaction) can be partly solved by developing a personality (or persona), created by programmers and manifesting through a specific linguistic behaviour as well as some meta-information available on the web site and/or on the profile of the chatbot in order to ‘justify’ its flaws (Lotze 2016, 39-40; Tewes 2005, 261).

Nowadays, chatbots play a more and more important role in the digital landscape. According to the IT security company Imperva Incapsula (ex Distil Networks), the non-human data traffic reached 37,9% of the entire data traffic on the Internet in 2019.<sup>3</sup> Although the first dialogue

---

<sup>3</sup> <https://www.imperva.com/resources/resource-library/reports/bad-bot-report-2019-the-bot-arms-race-continues/>.

systems were conceived in 1950 by the British mathematician Alan Turing and they were created starting from 1964,<sup>4</sup> chatbots began to attract more and more attention in the 1990s, with massive applications starting from the first years of the 21st century. Currently, chatbots are designed to solve several tasks according to their goals, regardless of whether they are commercial or non-commercial (Lotze 2016, 326; Storp 2002, 1-2). They even find an application in the area of general and language education. For example, the language learning platform Duolingo offered iOS users a set of three Duolingo Bots between 2016 and 2018 to practice respectively German, French, and Spanish through short chat-based dialogues (Mazzilli 2019, 145-7).

### 3 Related Research: Chatbot as Language Learning Tools

More and more often, the research has recently focused on chatbots as learning tools, although they are not just as present in the teaching practice, probably due to their critical aspects concerning structural and thematic incoherence (Bibauw, François, Desmet 2015, 60-2). In spite of these limits, the research has identified (also empirically) several applications and potentials of chatbots as learning tools and interlocutors in the L2,<sup>5</sup> e.g. high motivational potential, sensibilisation for several linguistic varieties etc. (Huang et al. 2017, 153-5; Wang, Petrina 2013, 127-8).

The scientific literature on the potential of chatbots for language education discusses their use in order to solve some common problems concerning language learning in formal contexts for adult learners, such as: (i) overcoming foreign language anxiety, which is meant as a response to “learning and using a language other than one’s mother tongue” (Tóth 2010, 18), (ii) developing communicative competence, defined by Balboni (2012, 72-3) as a combination of several partial competences involved in the interaction between mind and world, and (iii) dealing with conflicts in the L2, i.e. overcoming situations in which the communication fails due to misunderstandings, critical incidents etc. (Heringer 2017, 54-5). Some aspects of these problems can be partly solved by using digital learning instruments such as computers and smartphones as well as the corresponding forms of communication, e.g. chat (Grünewald 2016; Marques-Schäfer 2013, 42-6). Depending on their use, these means (including software) can

---

**4** The first chatbot ever developed was called ELIZA and was created by the computer scientist Joseph Weizenbaum in 1964.

**5** I focused especially on studies conducted in the area of English L2 or ESL, which offers a specific literature on the issue (Bibauw, François, Desmet 2015), and German L2, which is still developing in this direction (Wolski 2019; Mazzilli 2019).

be seen as (i) instruments generating learning contents and/or (ii) assistants guiding the learners throughout the learning process.

In particular, the research has focused on two general formats for the use of chatbots for foreign language education so far: (i) chatbots as interlocutors during 'free' dialogues, i.e. dialogues which are not characteristic for a specific situation dependent on place and time, and (ii) chatbots as interlocutors during the simulation of specific situations, i.e. for text-based roleplays concerning a specific setting (ticket booking, restaurant orders etc.). Although the first case can also be inherently a simulation (of human-human communication), it is not bound to a specific setting. Therefore, it differs from the second case of application.

Small talk pertains to the first category of applications (Goda et al. 2014, 1-3). Unlike interactions in the L2 with human interlocutors, interactions with chatbots are not bound to a specific setting (since chatbots are always available, as long as the Internet connection is stable). As non-human and non-judging interlocutors, chatbots tend to guarantee a stress-free and emotionally sustainable interaction, which can contribute to overcoming foreign language anxiety (Huang et al. 2017, 154-5; Rubesch 2013, 160-8).

In regard to the second category, unlike interactions with human interlocutors that are based on certain behavioural patterns, chatbots give unexpected answers, which the learners are confronted with in order to pursue their goal in a simulated, but still not easily predictable situation (Huang et al. 2017, 151-3; Istrate 2018, 471-3). Thus, the learners act in the L2 and use the L2 to enhance their communicative competence by interacting with chatbots that simulate specific roles and personas (waiter/tress, receptionist etc.).

Although chatbots and edu-bots are still interactionally rudimentary, the research has already identified and empirically examined several applications of chatbots for educational purposes, especially to solve the above-mentioned problems of language education for adults in an academic context. In particular, the following applications of chatbots contribute to the overcoming of foreign language anxiety and to the enhancement of communicative competence according to the research: (i) "impersonation game" (Carpenter, Freeman 2005, 1) to deal with cultural aspects of the L2, especially through the chatbot-based simulation of a specific personality, and (ii) chatbots as dialogue-based interactive glossaries (Bii 2013).

In all these cases, the research on the potential of chatbots for language education has led to partially contradictory results. On the one hand, chatbots consent a stress-free interaction in the L2 because of their motivating and non-human (i.e. non-judging) character (Goda et al. 2014, 2-4; Wang, Petrina 2013, 127); on the other hand, the realisation of a structurally and semantically coherent interaction is still problematic for the most chatbots (Lotze 2016, 381-3; Mazzilli 2018,

19-23). At least two elements have not been sufficiently considered yet, even in the most recent research on the topic: (i) the potential of chatbots (even the potential of their most critical aspects) as problematic interlocutors, and (ii) the learners' perception and reception of chatbots used for simulating and solving conflicts in the L2. These are the points of work for the present research.

## 4 Research Design

I offered my own contribution to the study of the impact of chat-based human-computer interaction on the acquisition and application of basic strategies to overcome conflicts in the L2 by designing and conducting the lesson/experiment *Konflikte im Alltag* ('conflicts in everyday life'). The main goals of the experiment were (i) to produce a concrete example for the use of chatbots as tools to enhance conflict-solving skills in the L2, (ii) to collect and analyse the corresponding empirical data gathered during and after the experiment, and (iii) to reflect on the use of non-educational chatbots for language learning as well as on how to enhance them and to what extent they can be used instead of or in combination with edu-bots in language education.

The experiment focused on the use of Elbot for German learning in an academic context. It consisted of a 75-minute lesson in the German language course for first-degree Foreign Languages students in their third and last year of study at the Aldo Moro University of Bari (Italy). This lesson was a single teaching event for the experimentation with Elbot for educational purposes. I held the lesson *Konflikte im Alltag* according to the principles of action-oriented learning (*handlungsorientierte Didaktik*, see Linthout 2004), which is based on the learners' holistic involvement, on the enhancement of their autonomy, and on the creation of a relaxed learning environment (Meyer 1996, 412-20; Linthout 2004, 7-8).

When the lesson/experiment was conducted, the 28 participants' German level was a B1, according to the Common European Framework of Reference for Languages.<sup>6</sup> They worked mostly in pairs during the lesson, which took place in a computer room at the above-mentioned university. Among the several learning goals, I focused on the following: (i) development of awareness in regard to the use of insults and hate speech on the Internet, (ii) better understanding of how conflicts work, (iii) adoption of basic strategies and use of appropriate expressions to solve conflicts, and (iv) development of digi-

---

<sup>6</sup> For further information, visit the following website: <https://www.coe.int/en/web/common-european-framework-reference-languages>.

tal literacy (i.e. the capability of dealing with technology in a critical way, see Marques-Schäfer 2013, 38), especially concerning chatbots.

During the lesson, the learners used Elbot to simulate and solve a conflict by intervening directly in the development of the open interaction<sup>7</sup> with Elbot, either individually or in pairs. The decision to use Elbot among many other educational and non-educational German-speaking alternatives is justified by the fact that the German version of this chatbot was developed to entertain German native speakers and not as an educational tool (Lotze 2016, 55-9). This means that Elbot offers authentic occasions to practice the language, according to the usual classification of learning materials in authenticity and non-authenticity categories. Elbot is also one of the most frequently examined chatbots in the German-speaking linguistic research, which makes it easier to interpret its interaction patterns and modes, e.g. irony. Besides, Elbot is free and provided with an explicative avatar [fig. 1] as well as a set of meta-information about its persona, e.g. its profession ('humanologist').

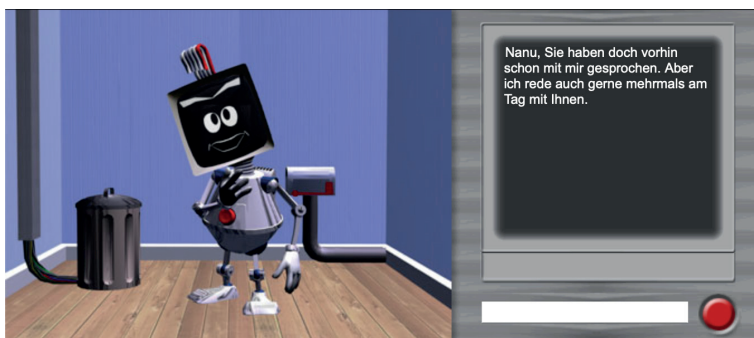


Figure 1 Elbot's chat interface. <http://www.elbot.de/htm/elbot/frame.htm>

The lesson and the corresponding worksheet consisted of eight activities, which refer to three different phases of the lesson: motivational/global phase (activities 1-3), elaboration/analysis (activities 4-7), and application/synthesis (activity 8). The learners used Elbot in the second phase of the lesson, which was also the longest one [tab. 1].

---

<sup>7</sup> The dialogues with Elbot were not constraint in any way, nor Elbot's responses were controlled during the lesson. The adequacy of Elbot's reactions in regard to the learning goals was guaranteed by the fact that, similarly to other chatbots, Elbot is programmed to temper any conflict and to keep the conversation going as long as possible, which results in a tendency to apply basic conflict-solving strategies every time the users send a potentially critical response, e.g. insults (Lotze 2016, 334-46).



**Table 1** Phases of the lesson/experiment *Konflikte im Alltag*

Phase	Duration	Description
Motivation (activities 1-3)	10 minutes	<ul style="list-style-type: none"> <li>- activation of previous knowledge about insults and hate speech in the L2 through visual inputs</li> <li>- comprehension and analysis of a web-based conflict in the L2 with focus on the use and on the effect of insults on the interaction</li> </ul>
Elaboration (activities 4-7)	50 minutes	<ul style="list-style-type: none"> <li>- familiarisation with Elbot through simple task-oriented dialogues</li> <li>- analysis of a running conflict with Elbot including reproductive and partially reproductive tasks with focus on conflict-solving strategies and expressions</li> </ul>
Application (activity 8)	15 minutes	<ul style="list-style-type: none"> <li>- autonomous production of a text</li> <li>- (Facebook comment) in order to solve a web-based conflict in the L2</li> <li>- possible contextualized use of expressions occurred during the previous activities</li> </ul>

For the creation of this communicative micro-training, I decided to profit from the human tendency to experiment and to use chatbots as a test stand, which is typical for bot talk<sup>8</sup> (Fischer 2010, 2350-2; Lotze 2016, 340-4). During the lesson, the learners reflected on the consequences of the use of hate speech and insults and classified Elbot's responses according to their impact on the conflict. Elbot's messages constituted a constantly growing glossary-like collection of useful expressions that the learners could use to reflect on the most adequate strategies to solve the conflict depending on the characteristics of the conflict itself. The activities 5, 6, and 7 (see below) were particularly relevant in this sense, since they were designed to stimulate the learners to participate in a conflict within the boundaries of human-computer interaction, which could protect them from stress. These activities followed some comprehension and analysis tasks (activities 1-3) as well as a short introducing dialogue between Elbot and every learner (activity 4). Activities 5-8 are listed below (for the original German version, see footnotes).<sup>9</sup>

<sup>8</sup> Bot talk (or computer talk) consists of the "several instances of deviant or odd formulations that look as if they were intended to be particularly suitable to use with a computer as the partner of communication" (Zoepritz 1985, 1). In particular, the term *bot talk* underlines the impact of software architecture and interface on the user's behaviour during the interaction (Mazzilli 2018, 18-19).

<sup>9</sup> Original German worksheet extract: "5. Versuchen Sie, den Chatbot mit verschiedenen Beleidigungen zu provozieren, und tragen Sie dessen Reaktionen je nach Ihrer Wahrnehmung in die Tabelle ein. Benutzen Sie dabei die Ausdrücke, die Sie in der Übung 1 gesammelt haben. // Positive Reaktion / Neutrale Reaktion / Negative Reaktion"

5. Try to provoke Elbot by using various insults and copy Elbot's reactions in the following table according to your perception. Please use the expressions collected in activity 1.

Positive reaction	Neutral reaction	Negative reaction

6. Analyse Elbot's reactions (see table in activity 5) and answer the questions below. Discuss in pairs.

- a) Which reactions do you find effective in order to create a win-win situation? Why?
- b) Has Elbot reacted ironically? If so, why?
- c) Which expressions used by Elbot are known/new to you? [...]

7a) Go on chatting with Elbot. Try to bring the interaction back to a calm, non-aggressive level by attempting several kinds of apologies. Then take notes of Elbot's response.

For example: *I made a mistake / I am sorry / ...*

Elbot's reaction: [...]

- b) Do you think that the conflict with Elbot is solved now? Why? Discuss in pairs. [...]

8. Roleplay. Go back to the Facebook discussion from activity 8. Imagine that you are the administrator of the Facebook group where the post was published. How would you as an admin solve the conflict between User 1 and User 2? Try writing a Facebook comment and compare then your comments in pairs.

(Excerpt from the worksheet used during the lesson/experiment *Konflikte im Alltag*)

Even though the (limited and controlled) use of insults and hate speech might be easily seen as questionable, it was always justified by and aimed to the analysis of Elbot's reactions. Like many other chatbots, Elbot is also programmed to react to any provocation in a conciliatory and non-aggressive way or sometimes ironically in order to temper the conflict (Fischer 2010, 2350-2; Lotze 2016, 340-4).

---

// 6. Analysieren Sie die Reaktionen von Elbot (s. Tabelle aus Übung 5) und beantworten Sie die Fragen. Diskutieren Sie zu zweit. // a) Welche Reaktionen finden Sie besonders effektiv, um eine Win-Win-Situation zu fördern? Warum? // b) Reagiert Elbot auch ironisch? Wenn ja, wann? // c) Welche Ausdrücke waren Ihnen bereits bekannt und welche sind neu? [...] // 7. a) Chatten Sie weiter mit Elbot. Versuchen Sie, die Interaktion wieder auf ein sachliches, nicht-aggressives Niveau zurückzuführen, indem Sie verschiedene Formen der Entschuldigung ausprobieren. Notieren Sie die Reaktion von Elbot. // Zum Beispiel: *Ich habe einen Fehler gemacht. / Es tut mir leid. / ...* // Reaktion von Elbot: [...] // b) Denken Sie, dass der Konflikt mit Elbot jetzt gelöst ist? Warum? Diskutieren Sie zu zweit. [...] 8. Rollenspiel. Gehen Sie zurück zur Facebook-Diskussion aus Übung 2. Stellen Sie sich vor, dass Sie VerwalterIn (Admin) der Facebook-Gruppe wären, in der der Beitrag gepostet wurde. Wie würden Sie als Admin den Konflikt zwischen Nutzer 1 und Nutzer 2 lösen? Schreiben Sie unten einen Facebook-Kommentar und vergleichen Sie dann Ihren Text mit dem Kommentar Ihres/r Nachbarn/in".

Thus, during the dialogue Elbot offers a wide range of useful expressions to solve several kinds of conflicts in German L2, which the participants could use for activity 8 (i.e. solving a simulated conflict between two human interlocutors). Through the interaction with Elbot, the learners experienced conflicts not as a given fact, but rather as a 'live' event in which they could intervene in order to (i) experiment several strategies to solve the conflict, and (ii) use the corresponding expressions in the L2 during a still running interaction with an interlocutor capable of giving unexpected responses.

## 5 Method

The experiment *Konflikte im Alltag* is part of a case study focused on two main objectives: (i) identifying how the learners perceive Elbot as an educational instrument, and (ii) defining to what extent Elbot influences the learners' development of basic strategies to manage conflicts. To pursue these two purposes, I designed this lesson/experiment also as a chance for gathering empirical data, which was collected, partly elicited and analysed according to a mixed method integrating (i) quantitative and qualitative analysis of the learners' responses to a questionnaire and (ii) qualitative and contrastive analysis of selected texts and notes written by the learners in their worksheets during the lesson while and after interacting with Elbot and solving the eight tasks.

Due to the specific character of this study, the number of participants (28) was considered adequate to collect some relevant data in regard to the research questions. In fact, this case study has three levels of specificity: (i) specific target and setting (first-grade students of German L2 in the field of intercultural studies and modern philology), (ii) specific chatbot as a language learning tool (German-speaking and non-educational chatbot Elbot), and (iii) specific learning goals (analysis of conflicts in the L2 and acquisition as well as application of expressions and strategies to overcome conflicts in the L2).

The data was collected and elicited respectively during and after the lesson, which was briefly described in the previous section. Before the lesson, the participants were informed about the types of data needed as well as the non-human nature of the interlocutor they were going to work with [tab. 2]. In so doing, I made sure that the learners were aware of the potential benefits and risks of interacting with a chatbot, especially in a language (German) in which they were still on an intermediate level of knowledge at the time of the lesson.

**Table 2** Phases of the experiment

Phase of the experiment	Duration (minutes)	Activities
Preparation	10	<ul style="list-style-type: none"> <li>- registration of the participants</li> <li>- clarification of goals and non-human nature of Elbot</li> <li>- distribution of worksheets</li> </ul>
Lesson	75	<ul style="list-style-type: none"> <li>- focus on the eight worksheet tasks</li> <li>- task-oriented chat with Elbot</li> <li>- discussion and work in pairs</li> </ul>
Conclusion	15	<ul style="list-style-type: none"> <li>- distribution and completion of questionnaires</li> <li>- submission of questionnaires and worksheets</li> </ul>

## 5.1 Participants

At the time of the experiment, all 28 participants attended the third-year curricular German course (target level: B2) during their first-level degree at the Faculty of Foreign Languages and Literatures at the Aldo Moro University of Bari. No participant was recruited outside of this course. This is justified by the intention to work with an already integrated group generally used to and capable of working together. This aspect is crucial, since most tasks needed to be solved in pairs and the learning atmosphere had to stay as stress-free as possible, according to the principles of action-oriented learning (Linthout 2004, 7-8; Meyer 1996, 412-20).

Before starting the lesson, I made sure that the participants were adequately informed about the nature of the interaction and expressed their informed consent, which each learner did. During the completion of the questionnaires (see below), the participants outlined a rough profile of themselves. According to their answers, 27 out of 28 learners were 25 years old or younger at the time of the lesson. Besides, 57,1% of the learners studied Intercultural Communication and International Relations at the time of the lesson, whereas the remaining students' studies focused on Tourism Studies and Modern Languages and Cultures.

In regard to their experience with chat-based human-computer interaction, 67,9% of the learners declared in the same questionnaire that they had never interacted with a chatbot, whereas 60,7% of the group had already heard of chatbots [fig. 2]. Besides, the same amount of participants was not sure if they would have used a chatbot after the experience with Elbot, which could be linked to the fact that this was their very first experience with a chatbot for most of them. Anyway, this assumption will be discussed below.

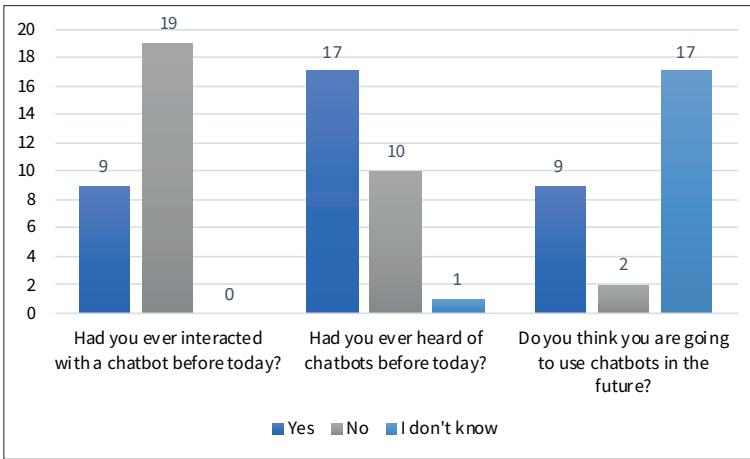


Figure 2 Participants' experience with chatbots

## 5.2 Data Elicitation and Analysis: Focus on Learners' Perception

In order to determine the learners' perception of the interaction with Elbot during the micro-training, I proposed an anonymous questionnaire in five parts, which I analysed both on a quantitative and qualitative level. In particular, the questionnaire contained totally 29 items and was produced in the learners' native language (i.e. Italian). It consisted of following sections: (i) personal data (age, course of studies, self-evaluated German level), (ii) previous experience with chat-based human-computer interaction, (iii) evaluation of the interaction with Elbot (technical issues, incoherence and quasi-coherence etc.), (iv) evaluation of Elbot's role during the lesson (quality of contribution, relevance in regard to the learning goals), and (v) free comments and observations about Elbot as a language learning tool.

The questionnaire contained both closed questions (producing nominal and categorical data) and open questions. These were added in order to allow the participants to motivate, explain, and specify their answers. The resulting data was manually digitalized to facilitate the analysis. Even the very few cases of missing answers were appropriately signalled. Responses to open questions were analysed in combination with the corresponding closed question (if available) and used for investigating the learners' guided evaluation of Elbot for enhancing their conflict-solving skills.

### 5.3 Data Collection and Analysis: Focus on Elbot's Influence on Human Behaviour

In order to examine the impact of Elbot on the learners' linguistic behaviour during and after the interaction with the chatbot, I focused on their lexical choices during a separate activity that did not involve the use of Elbot. This included the collection and contrastive analysis of two kinds of texts (both in German): (i) Elbot's responses noted by the learners and classified as positive, neutral or negative reactions during activity 5 (see above) and (ii) learners' simulated Facebook comments in which they acted as administrators of a Facebook group in order to solve a conflict between two group members (activity 8, see above).

After the experiment, I digitalized also both notes and simulated Facebook comments. I chose not to correct the grammar errors occurring in some texts in order to offer a faithful overview of the learners' choices throughout the writing process. I cleaned up the short corpus resulting from the data collection by ruling out all possibly partial and/or unreadable contributions. At the end of this process, I decided to focus the qualitative analysis of the texts on nine worksheet samples, i.e. 18 sets of data, due to the elaborateness and explicative value of the conflict-solving strategies they contained.<sup>10</sup>

For each worksheet sample, I compared the learners' notes and simulated Facebook comments focusing on: (i) whether learners used any expressions attributed to Elbot during activity 5 in another context, (ii) what expressions the learners chose to reuse, (iii) why they decided to reuse a specific expression attributed to Elbot, and (iv) how the reused expressions were (correctly or incorrectly) varied and adapted to the new context. In so doing, I intended to show if and how experiencing a conflict with Elbot and noting its conciliating responses influenced the learners' strategies and lexical choices while trying to solve a conflict in a different context.

---

**10** Although the amount of analysed worksheets may seem limited (9 out of 28), this choice is due to the priority assigned to the depth of the qualitative analysis in this part of my research rather than its quantitative representativity, which is rather covered by the examination of the questionnaires (see above and Page et. al 2014, 85-93).

## 6 Selected Results

### 6.1 Analysis and Discussion: Focus on Learners' Perception

In the following paragraphs, I will focus on the data resulting from the learners' responses to the fourth section of the questionnaire, concerning the perception of Elbot's role during the lesson. In particular, the questions pertain to the interactions with Elbot that the learners engaged in order to solve tasks 5-7 (see above). The present analysis focuses on following topics: (i) Elbot's relevance during the lesson, (ii) technical and interactional problems, (iii) motivation through chat-based human-computer interaction, (iv) simulation of a conflict with human interlocutors vs. simulation of a conflict with a chatbot, (v) consistency with teaching method and learning goals, and (vi) authenticity of the interaction.

According to the results of the quantitative and qualitative analysis of the questionnaires, the learners perceived the interaction with Elbot as barely authentic, but also as less stressful and more motivating than human-human communication in the L2. According to 53,6% of the participants, Elbot's responses were only partly relevant in regard to the learning goals [fig. 3]. As the participants observed answering the corresponding open question, the dubious relevance of Elbot for pursuing the learning goals is due to the fact that (i) the interactions with Elbot contributed to reaching other goals which did not correspond to the planned learning goals (e.g. expanding vocabulary in other areas, such as, reportedly, science and technology), and (ii) the coherence problems of the chatbot sometimes distracted the learners from their actual goals.

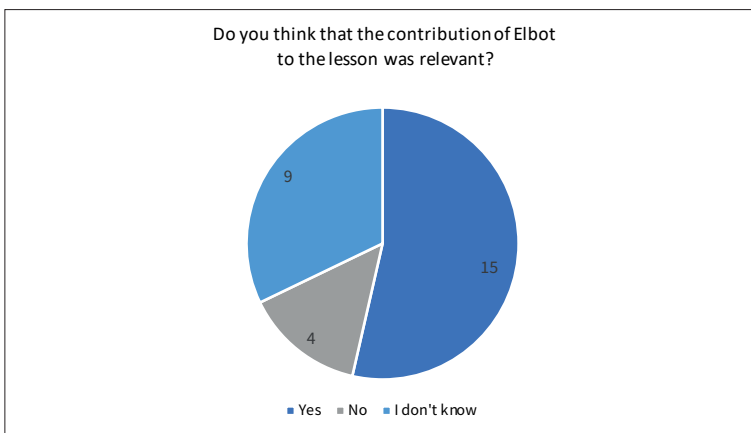


Figure 3 Learners' perception of Elbot's relevance

Therefore, Elbot's relevance was not entirely reduced by the partial inadequacy of the chatbot as an interlocutor, but rather by the simultaneous presenting of further and unexpected learning occasions during the conversation, suggesting that a more appropriate goal/task balance could be a crucial premise for the use of non-educational chatbots as language learning tools. Besides, the occurrence of unexpected inputs is also an important part of every learner-teacher interaction focused on the importance of the human component in the learning process, regardless of how formal and planned the lesson is.

Only 2 out of 28 learners reported a technical problem. In response to the corresponding open question, they learners stated that the coherence problems of the chatbot were slightly demotivating but still easy to deal with by applying specific strategies that the learners pursued during the conversation. Besides, the majority of the participants perceived the experience of task-oriented chatting with Elbot as motivating (60,7%) [fig. 4] and felt more confident while dealing with conflicts with the chatbot than while facing them with humans (75%) [fig. 5], which partly confirmed the results of the research on the topic so far (Huang et al. 2017, 153-4; Wang, Petrina 2013, 127-9). In fact, according to 75% of the learners, using hate speech with a human would have discomfited them, even in a controlled and simulated conflict. Knowing that the chatbot would not 'judge' them, the participants felt free to experiment with the L2 without worrying about their 'image'. Nevertheless, this result must be verified through further experiments ideally focused on a comparison between the learners' reaction to conflicts with human and non-human interlocutors, which was not the main goal of my research and has not been investigated in this study.

The relevance of stress-free experimentation with language and communication is particularly evident in the light of learning methods such as 'try and error' and 'learning by doing', which also constitute the main pillars of the above-mentioned action-oriented learning. Nevertheless, the credibility of Elbot as an interlocutor is still a problematic aspect: due to the coherence issues occurred during the interaction, the participants were still unsure about the authenticity of the experience. Besides, in spite of their curiosity towards the chatbot and human-computer interaction in general, some participants replied the open questions in this section by defining Elbot's responses as difficult to understand, since they were either bound to a very specific semantic area (especially technology) or syntactically and lexically on a higher level than the participants' B1, with a particular reference to Elbot's use of irony, which was not always detected by the learners. In this case, the fact that the interactions with Elbot were embedded in a wider formal context was crucial, since the learners had the chance to discuss Elbot's responses, research autonomously or ask for help in case of doubt. In this sense, the combination of



the use of a chatbot and the general teaching method applied to the whole lesson seemed to work for the 67,9% of the learners [fig. 4].

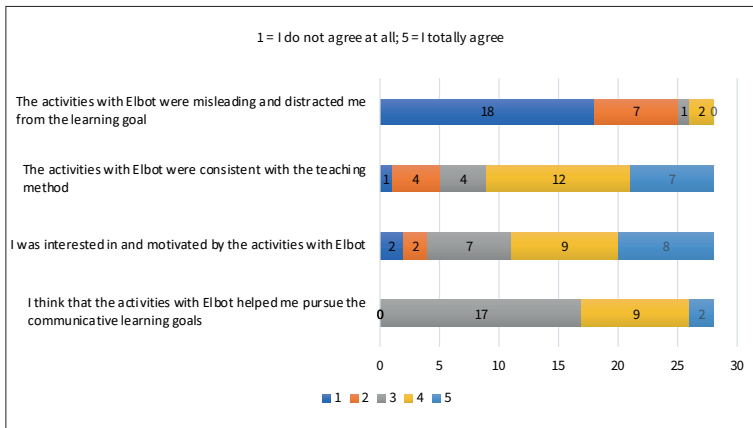


Figure 4 Learners' perception of the interactions with Elbot (1)

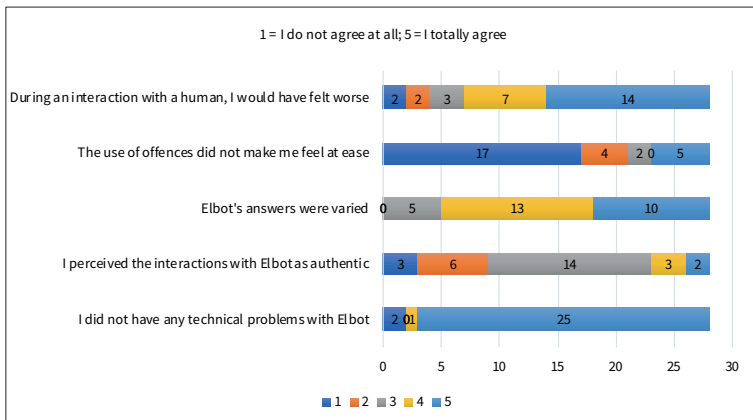


Figure 5 Learners' perception of the interactions with Elbot (2)

In conclusion, apart from the sometimes dubious credibility of the interaction, the learners perceived the activities with Elbot and Elbot's contribution itself as very useful, especially in a formal context. Over 90% of the learners stated that the activities involving an interaction with Elbot helped them pursue their communicative learning goals, such as the development and application of strategies to deal with conflicts in the L2.

## 6.2 Analysis and Discussion: Focus on Elbot's Influence on Human Behaviour

The contrastive analysis of the noted responses of the chatbot (activity 5) and the learners' texts (activity 8) also gave some interesting insights. In some cases, the learners' texts contained syntagms or even entire clauses attributed to Elbot (so-called recurrences). During the qualitative analysis of the samples, I identified three levels of recurrence, each one corresponding to three samples:<sup>11</sup> (i) zero recurrence, (ii) partial recurrence, and (iii) extended recurrence. In the first case, no similarities between Elbot's responses and learners' texts were found. In all the samples pertaining to this category, Elbot's responses which were noted by the learners were either syntactically and lexically complex (Lotze 2016, 322-34), if not incomprehensible to the learners, or thematically chatbot-specific (e.g. they contained frequent references to the non-human nature of the speaker, such as 'we robots'), or they even pursued a conflict-solving strategy which was different from the one chosen by the learners. For these reasons, the learners did not use the noted responses of the chatbot as a source for their own conflict-solving comment. It is still unclear if they made this decision consciously or on a pre-conscious level.

The other two levels are much more interesting. Some of the learners' texts showed an evident influence of the interaction with Elbot, be that expressed through the recurrence of single words, syntagms or entire clauses previously used by Elbot. For example, in the case of partial recurrence, one participant (P<sub>1</sub>) reused a syntagm consisting of article-adjective-noun which he/she had noted as one of Elbot's positive reactions during activity 5:

If I were in your place, I would be careful. One word from me and you get deleted from all computers in the world. But if you apologize, then I will put in a *good word* [for you] with my cousin, who examines tax returns. (Elbot's response as noted by P<sub>1</sub> in activity 5)

People, you have two different opinions, it doesn't make sense to offend other people. You can simply talk, use a *good word* and compare your opinions and if you don't find a solution, no problem! It's not the end of the world! (Simulated Facebook comment written by P<sub>1</sub> in activity 8)<sup>12</sup>

**11** Functional words such as articles and prepositions were excluded from the analysis, since their recurrence was not considered relevant or indicative of a relation between the interaction with Elbot and the writing process in activity 8.

**12** All italics in the original texts and in the translations are mine. Original notes and text (German): "An Ihrer Stelle wäre ich da ein wenig vorsichtig. Ein Wort von mir und Sie werden aus sämtlichen Computern der Welt gelöscht. Aber wenn Sie sich entschuldigen"

As this example shows, the recurring syntagms were not always used with the same meaning as the source note. Sometimes, the learners interpreted and used them in a different way. In this case, P<sub>1</sub> broke up the phrase *ein gutes Wort einlegen* ('to put in a good word') as reported in the noted response by Elbot by using only the nominal syntagm *ein gutes Wort* ('a good word'). This suggests that P<sub>1</sub> could have not fully understood the meaning of the original phrase in the source note, which would be compatible with the above-mentioned learners' comments on the syntactical and lexical complexity of Elbot's responses, and he/she decided to use only a part of the phrase in his/her text without adapting it to the new context of the target text.

In other cases, entire clauses attributed to Elbot in activity 5 were reused by the learners in their own texts, either unvaried or in a slightly modified form. When they were modified, the clauses were adapted at least thematically and functionally to the new context of the Facebook comment. This adaptation was not always followed by the necessary syntactic adjustments, which showed on the one hand that the learners managed to make sense of their interaction with Elbot as well as of its responses, and on the other hand that they were still not able to adapt the reused clauses to a relatively complex syntax in the target text. For example, one participant (P<sub>2</sub>) reused an entire clause previously attributed to Elbot by (i) deleting and substituting the words that did not suit the new context of the Facebook comment on a semantic level, e.g. "arme kleine Roboter" ('poor little robots', see below), and (ii) leaving the syntax of the sentence unchanged. In so doing, P<sub>2</sub> did not follow the German grammar rule establishing that the conjugated verb must be put at the very end of a declarative clause starting with *dass* ('that'), which is a so-called *Dass-Satz* (literally 'that-clause').

I hope you can do better than offending poor little robots. (Elbot's response as noted by P<sub>2</sub> in activity 5)

Hello User 1 and User 2, our admin will analyse your offer. We hope that this kind of behaviour will not be reiterated in the future and that you can do better than offending others. (Simulated Facebook comment written by P<sub>2</sub> in activity 8)<sup>13</sup>

---

gen, lege ich *ein gutes Wort* bei meinem Vetter ein, der Steuererklärungen prüft. // Leute, / ihr habt zwei verschiedenen Meinungen, es hat kein Sinn die anderen zu beleidigen. Ihr könnt einfach sprechen, *ein gutes Wort* benutzen und die Meinungen vergleichen und wenn ihr keine Solution finden, kein Problem! Das ist doch nicht das Ende der Welt!"

**13** Original note and text (German): "Ich hoffe, sie können mehr als arme kleine Roboter zu beleidigen // Hallo Nutzer 1 und 2, unsere Admin wird seine Angebot analysiert. Wir hoffe, dass diese Verhaltensweisen nicht meher in der Zukunft wiederholt werden, und dass Sie können mehr als andere zu beleidigen".

In conclusion, it is clear that the recurrence of any linguistic elements previously used by Elbot (syntagms or entire clauses) depends on the syntactic and thematic complexity as well as the usability of Elbot's responses in other contexts and/or for the adoption of other strategies to solve the conflict. In case of partial or extended recurrence, these phenomena do not always imply the adaptation of the reused linguistic elements to the target text and they do not necessarily indicate that the learners understood and actively elaborated the recurrent elements.

## 7 Conclusion

My research on the potential of chatbots as tools for enhancing conflict-solving skills in the L2 in a formal context led to the result that chatbots (in this case Elbot) are particularly useful for the simulation of conflicts of any kind which the learners try to solve by experimenting with different strategies. Elbot contributed to a stress-free and motivating learning atmosphere, since its non-human nature encouraged the learners to feel free to experiment with the L2 without being 'judged' by the interlocutor. The necessary reflection on their choices resulted rather from the activities that engaged the learners in pairs as well as the entire group (including the teacher) in a concomitant meta-communicative process. In this way, the learners could intervene in the running conflict in real time to verify and reflect on the validity of each strategy. In so doing, one of the most questionable characteristics of bot talk (i.e. the use of hate speech) and the tendency of chatbots to conciliate in order to keep the interaction going were fully used to develop the learners' skills in the L2.

On the other hand, the learners' perception of Elbot's credibility as an interlocutor as well as their mostly non-adapted reuse of Elbot's responses showed how crucial the human component is in language education. This includes not only the entire group of learners, but also the teacher, who should support the learners' autonomous confrontation with the chatbot and actively participate in the reflection process.

Based on these results, I conclude that (i) even the problematic aspects of chat-based human-computer interaction have a potential for language education, since the chatbot as a non-native speaker of the natural language (Lotze 2016, 167-8) can help simulate conflicts that belong to everyday life (especially misunderstandings and critical incidents), and (ii) the chatbot participates in the didactic triangle as an instrument and not as an interlocutor, since it is still not perceived as a credible chat partner by the learners. This implies that it is possible to use existing non-educational chatbots for language education, although (or even given the fact that) they are not technically flawless.

Nevertheless, this *does not* imply that the research and the development of chatbots should not deal with any technical issue. Rather,

these results suggest that the research should focus not only on the technical innovations, but also on valid strategies to use existing chatbots as authentic tools, i.e. as tools which were originally developed for native speakers of the L2, and to create edu-bots for language learning in a blended learning framework. For example, software developers should cooperate with language educators in order to offer an educational version of existing chatbots, which has following advantages: (i) the database would contain authentic language material (i.e. designed for native speakers of the L2), (ii) the adequate feedback offered by the chatbot would be based on the language use attested by the content of the database rather than on uncontextualized grammar rules, and (iii) the learners would have the chance to change from the educational to the non-educational version of the chatbot based on their actual needs and tasks (need for analytic tasks and feedback vs. need for 'casual' conversation and small talk in the L2).

For example, an educational version of Elbot could give the users an understandable feedback regarding the thematic adequacy and grammatical correctness of their responses by 'dramatizing' the feedback through the specific characteristics of its persona (i.e. a 'humanologist' that examines human beings and their behaviour in several simulated situations). In this way, the users would experience the phenomenon of making mistakes (which is necessary for language learning) with an already known tool which is non-judging, but also capable of offering an adequate feedback, i.e. a feedback that allows to negotiate meanings and structures without suppressing the experience of error.

## References

- Balboni, P.E. (2012). *Le sfide di Babele: Insegnare le lingue nelle società complesse*. Torino: UTET Università.
- Bibauw, S.; François, T.; Desmet, P. (2015). "Dialogue-Based CALL: An Overview of Existing Research". Bradley, L.; Thouësny, S.; Helm, F. (eds), *Critical CALL = Proceedings of the 2015 EUROCALL Conference, Padova, Italy*. EUROCALL. Voillans: Research-publishing.net, 57-64.
- Bii, P. (2013). "Chatbot Technology: A Possible Means of Unlocking Student Potential to Learn How to Learn". *Educational Research*, 4(2), 218-21.
- Carpenter, R.; Freeman, J. (2005). "Computing Machinery and the Individual: the Personal Turing Test". <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.4547&rep=rep1&type=pdf>.
- Fischer, K. (2010). "Why It is Interesting to Investigate How People Talk to Computers and Robots". *Journal of Pragmatics*, 42, 2349-54.
- Goda, Y. et al. (2014). "Conversation with a Chatbot Before an Online EFL Group Discussion and the Effects on Critical Thinking". *The Journal of Information and Systems in Education*, 13(1), 1-7.
- Grünewald, A. (2016). "Digitale Medien und soziale Netzwerke im Kontext des Lernens und Lehrens von Sprachen". Burwitz-Melzer, E. et al. (Hrsgg), *Handbuch Fremdsprachenunterricht*. Stuttgart: UTB, 463-6.

- Heringer, H.J. (2017). *Interkulturelle Kommunikation*. Stuttgart: UTB.
- Huang, J.-X. et al. (2017). "A Chatbot for a Dialogue-Based Second Language Learning System". Borthwick, K.; Bradley, L.; Thouëсны, S. (eds), *CALL in a Climate of Change: Adapting to Turbulent Global Conditions*. Voillans: Research-publishing.net, 151-6.
- Istrate, A.-M. (2018). "Artificial Intelligence and Machine Learning-Future Trends in Teaching ESL and ESP". *eLearning & Software for Education*, 2, 471-6.
- Kansanen, P.; Meri, M. (1999). "The Didactic Relation in the Teaching-Studying-Learning Process". *Didaktik/Fachdidaktik as Science(-s) of the Teaching profession*, 2(1), 107-16.
- Kerly, A.; Hall, P.; Bull, S. (2006). "Bringing Chatbots into Education: Towards Natural Language Negotiation of Open Learner Models". *Knowledge-Based Systems*, 20(2), 177-88. [https://doi.org/10.1007/978-1-84628-666-7\\_14](https://doi.org/10.1007/978-1-84628-666-7_14).
- Linthout, G. (2004). *Handlungsorientierter Fremdsprachenunterricht: Ein Trainingsprogramm zur Kompetenzentwicklung für den Beruf*. Amsterdam: Rodopi.
- Lotze, N. (2016). *Chatbots. Eine linguistische Analyse*. Pieterlen; Bern: Peter Lang Verlag.
- Marques-Schäfer, G. (2013). *Deutsch lernen online*. Tübingen: Gunter Narr Verlag.
- Mazzilli, F. (2018). "Bot talk e apprendimento linguistico. L'uso dei chatbot per lo sviluppo della competenza comunicativa nella lingua straniera". *Iperstoria*, 12, 13-36. <https://doi.org/10.13136/2281-4582/2018.i12.387>.
- Mazzilli, F. (2019). "Kriterien und Vorschläge zur Evaluation textbasierter Chatbots für die touristische Sprachbildung am Beispiel von Lufthansa BestPrice und Duolingo Bots". *Sprache im Beruf*, 2(2), 142-64.
- Meyer, H. (1996). *Unterrichtsmethoden I: Theorieband*. Frankfurt/Main: Cornelsen Scriptor.
- Page, R. et al. (2014). *Researching Language and Social Media: A Student Guide*. London: Routledge.
- Rubesch, T. (2013). "Interactive Competence in Student Use of a Conversational Agent". *Studies in Linguistics and Language Teaching*, 24, 157-72.
- Russell, S.J.; Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. London: Pearson Education Limited.
- Storp, M. (2002). "Chatbots. Möglichkeiten und Grenzen der maschinellen Verarbeitung natürlicher Sprache". *Networx*, 25, 1-34.
- Tewes, M. (2005). "'Sie können ruhig in ganzen Sätzen mit mir sprechen!' Chatbots und ihre Bedeutung für internetbasierte Kommunikation". Siever, T.; Schlobinski, P.; Runkehr, J. (Hrsgg), *Websprache.net: Sprache und Kommunikation im Internet*. Berlin; New York: Walter de Gruyter, 242-65.
- Tóth, Z. (2010). *Foreign Language Anxiety and the Advanced Language Learner: A Study of Hungarian Students of English as a Foreign Language*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Wagner, C. et al. (2012). "When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks". *Making Sense of Microposts*, 2(4), 41-8.
- Wang, Y.F.; Petrina, S. (2013). "Using Learning Analytics to Understand the Design of an Intelligent Language Tutor – Chatbot Lucy". *International Journal of Advanced Computer Science and Applications*, 4(11), 124-31. <https://doi.org/10.14569/ijacsa.2013.041117>.
- Wolski, P. (2019). "Förderung der berufsbezogenen Gesprächskompetenz durch interaktive Kommunikationssysteme – theoretische Grundlagen und praktische Anwendungen". *Sprache im Beruf*, 2(2), 133-41.
- Zoeppritz, M. (1985). *Computer Talk?* Technical report TN 85.05. Heidelberg: IBM Heidelberg Scientific Center.