



# Language Learning with Artificial Entities: Effects of an Artificial Tutor's Embodiment and Behavior on Users' Alignment and Evaluation

Astrid Rosenthal-von der Pütten<sup>1</sup>(✉) , Carolin Straßmann<sup>2</sup> ,  
and Nicole Krämer<sup>3</sup> 

<sup>1</sup> iTec - Chair for Individual and Technology,  
RWTH Aachen University, 52078 Aachen, Germany  
arvdp@humtec.rwth-aachen.de

<sup>2</sup> Institute of Computer Science, University of Applied Sciences Ruhr West,  
46236 Bottrop, Germany

carolin.strassmann@hs-ruhrwest.de

<sup>3</sup> Social Psychology: Media and Communication, University of Duisburg-Essen,  
47057 Duisburg, Germany

nicole.kraemer@uni-due.de

<https://www.itec.rwth-aachen.de>, <https://informatik.hs-ruhrwest.de>,  
<https://www.uni-due.de/sozialpsychologie/kraemer.shtml>

**Abstract.** Based on the assumption that humans align linguistically to their interlocutor, the present research investigates if linguistic alignment towards an artificial tutor can enhance language skills and which factors might drive this effect. A  $2 \times 2$  between-subjects design study examined the effect of an artificial tutor's embodiment (robot vs. virtual agent) and behavior (meaningful nonverbal behavior vs. idle behavior) on linguistic alignment, learning outcome and interaction perception. While embodiment and nonverbal behavior affects the perception of the tutor and the interaction with it, no effect on users' linguistic alignment was found nor an effect on users learning outcomes.

**Keywords:** Social robot · Virtual agent · Embodiment · Nonverbal behavior

## 1 Introduction

When native speakers and non-native speakers meet, native-speakers often adapt to non-natives in order to foster mutual understanding and successful communication, sometimes with the negative outcome of interfering with successful second language acquisition (SLA) on a native-speaker level. As a matter of course, native speakers do not mean to decrease learning progress, but rather engage in an automatic behavior of adaptation to their interlocutor. Using artificial tutors could help to overcome this bias as they can be designed to not or only

slightly adapt to human users thereby given a better example of correct speech. Moreover, the phenomenon of computer talk, i.e. users more strongly align to computers in order to ensure communicative success, could be exploited for SLA using the exaggerated alignment tendencies of users confronted with computers. Alignment is seen as core to language acquisition, thus, also to SLA [1] and the tendency of non-natives to align to technology in a learning setting could be exploited for SLA. Admittedly, system characteristics have to be taken into account and their potential inhibiting or facilitating effects need to be explored. Hence, the current work investigates if linguistic alignment towards an artificial tutor can enhance language skills and which factors might drive this effect.

### 1.1 Linguistic Alignment HCI and Second Language Acquisition

Whilst in conversation, interaction partners align linguistically on different levels, for instance, regarding accent or dialect [9], lexical choices and semantics [6] as well as syntax [4]. Similar tendencies have been observed in interactions with artificial interlocutors mirroring the effects of alignment with regard to prosody, lexis, and syntax (for an overview cf. [5]). Comparative studies indicate that people tend to show stronger alignment with computers presumably to compensate the computers weaker communicative abilities, a phenomenon known as computer talk [5]. However, “when social cues and presence as created by a virtual human come into play, automatic social reactions appear to override initial beliefs in shaping lexical alignment” [3] resulting in slight decreases of the computer talk phenomenon. A first study with native and non-native speakers showed that both groups aligned lexically to a virtual tutor [17]. In a precursor to the current study we explored whether alignment with an artificial tutor in a SLA setting improves language skills and whether this is influenced by the tutor’s embodiment (voice-only, virtual or physical embodiment) or type of speech output (text-to-speech or prerecorded language, cf. [16]). Although participants aligned to the artificial tutor in all conditions comparably to previous studies, the alignment was not correlated with post-test language skills. Moreover, the variation of system characteristics had barely influence on the evaluation of the system or participants’ alignment behavior, neither for embodiment nor for quality of speech output. The present study shall provide additional evidence on whether this finding is persistent and whether the specific behavior of the artificial tutor influences the results.

### 1.2 Effects of Differently Embodied Artificial Entities

Virtual agent or robot? This is an ongoing debate in the research community developing and evaluating embodied conversational agents. Both embodiment types provide unique interaction possibilities, but also come along with certain restrictions (for an overview cf. [10,14]). Indeed, studies comparing the two embodiment types have led to inconsistent results. While a majority of study

results supports the notion that robots are superior to virtual representations (e.g. perceived social presence [7], entertainment and enjoyment [12], trustworthiness [12], persuasiveness [13], and users’ task performance [2]), other results suggest virtual representations are more beneficial especially in conversational settings [13]. In fact, there seems to be an interaction effect of embodiment and task [11] suggesting that robots might be better suited in (hands-on) task-related scenarios, while virtual agents could be beneficial for purely conversational settings. This is probably due to the perception of different bodily-related capabilities of virtual agents and robots which might lead to different expectations for the subsequent interaction and thus also different outcomes as suggested by the EmCorp framework of Hoffmann et al. [10]. In this regard the displayed nonverbal behavior plays an important role. For instance, it was suggested that if an entity has the capability for nonverbal behavior, but does not use, for instance, gestures in situations where they would be beneficial this leads to negative evaluations [10]. Only two studies have looked into the effect of different embodiment types on participants’ linguistic behavior. Fischer [8] found that verbosity and complexity of linguistic utterances did not differ between a virtual agent or a robot, but participants used more interactional features of language towards the robot such as directly addressing it by its name. The interplay of embodiment and linguistic alignment has been investigated in our prior work [16] where we found no differences in users’ evaluation of and alignment towards the different version of the tutor. However, in this study we did not explicitly address the nonverbal behavior of the artificial tutor and accordingly the perception of bodily-related or communicative capabilities of the entities which will be manipulated in the present work.

### 1.3 Research Questions and Hypotheses

In this work, the two central questions we examine are whether an artificial tutor’s embodiment (virtual agent vs. robot) influences participants’ evaluation of the tutor, their lexical and syntactical alignment during interaction and their learning effect after the interaction (RQ1) and whether displayed nonverbal behavior will affect users’ alignment (RQ2). Since previous work showed that robots can elicit more positive evaluations than virtual agents [10, 14], we additionally propose that the robot will be rated more positively than the virtual agent (H1). Based on the EmCorp Framework by Hoffmann et al. [10], evaluations regarding the perceived bodily-related and communicative capabilities should differ (H2). Moreover, we investigate the effect of expressive nonverbal behavior on evaluation assuming that gestures lead to more positive evaluations (H3).

## 2 Method

### 2.1 Experimental Design and Independent Variables

In order to address our research questions we used a  $2 \times 2$  between-subjects design with embodiment and nonverbal behavior as independent variables.

Regarding the embodiment, participants either interacted with the physically embodied Nao robot or with a virtual version of the Nao robot (cf. Fig. 1). Secondly, we varied whether the artificial tutor exhibited meaningful nonverbal behavior (usage of deictic, iconic, and beat gestures and socio-emotional gestures) or only displayed idle behavior (very subtle changes in head position to show the robot is somewhat “alive”).



**Fig. 1.** Participant playing the guessing game with the virtual Nao

## 2.2 Participants and Procedure

Eighty-five volunteers (33 female, 52 male) aged between 18 and 40 years ( $M = 24.81$ ,  $SD = 3.84$ ) participated in this study. Participants stem from more than 35 different countries, speak more than 24 different native languages and exposed different levels of German language skills (minimum of an intermediate level). Participants were recruited on campus or in German classes in the local adult education center. The study was approved by the local ethics committee. The procedure, game materials and coding procedures for verbal behaviors are identical to those used in our prior work [16]. Upon arrival to the first study session participants read and signed informed consent. They completed two language tests: a test on grammar and reading and listening comprehension and a so called C-Test ([www.c-test.de](http://www.c-test.de)), a cloze test which addresses language skills with regard to different dimensions. Finally, they were invited for a second appointment. Based on their test results, their country of origin and first language, respectively, participants were distributed equally across conditions where possible for the second session. On the second appointment participants were instructed about the different tasks to be solved with the artificial tutor.

Each task was again explained by the tutor during the interaction (cf. Fig. 1). Participants were also given a folder with detailed instructions in case they did not understand the tutor. Participants completed five tasks: 1) introducing themselves, 2) describing a picture in detail, 3) playing a guessing game, 4) playing a search game, and 5) again describing a picture. The order of tasks was always the same for all participants. The first two tasks were used to make participants comfortable at speaking loudly to the system. The two structured games (guessing game and search game) were used to analyze alignment processes. In order to create a more believable training environment for the participants, we repeated the task of describing a picture to give participants another possibility to speak quite freely at the end of the learning session. After the interaction, participants completed a second C-Test as a measure of learning outcome and a questionnaire asking for their experiences and assessment of the interaction. Finally, they were debriefed, reimbursed (€10) and thanked for participation.

### 2.3 Dependent Variables: Self Report

For all items in all scales participants gave ratings on a 5-point Likert scale ranging from “I do not agree” to “I agree”.

**Person Perception.** As for the person perception of the robot, we measured *Likability* with a scale of eight ad-hoc generated items (e.g., The robot is friendly, likable, pleasant, warm ) which showed good internal consistency, Cronbach’s  $\alpha = .824$ . The robots perceived *Competence* was measured with eight ad-hoc generated items (e.g., The robot is intelligent, competent ; Cronbach’s  $\alpha = .751$ ). Lastly, we measured perceived *Autonomy* of the robot with five ad-hoc generated items. One item (The robot is not autonomous) seemed to have caused rating artefacts because of the negation and thus was deleted from the scale to increase internal consistency. The remaining four items in the scale showed good internal consistency (The robot is self-dependent, free, self-determined, responsible for its actions; Cronbach’s  $\alpha = .810$ ).

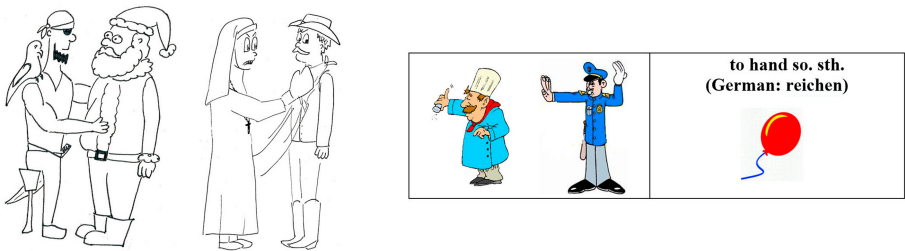
**Bodily-Related Capabilities.** In order to measure whether the physical or virtual embodiment of the artificial tutor results in different perceptions of their bodily-related capabilities we used 17 ad-hoc generated items that covered three factors: *Presence, Mobility and Tactile Interaction* and *(Shared) Perception*. This ad-hoc generated scale was actually a precursor of the EmCorp Scale [10]. Perceived presence of the robot was assessed using three items and showed sufficient internal reliability (The robot was in the same room with me; The robot was very present; The robot was not really present; Cronbach’s  $\alpha = .716$ ). We used four items to measure perceptions of Mobility and Tactile Interaction which also showed good internal consistency (The robot is able to move around in the room; The robot is able to walk towards me; The robot would have been able to touch me at any time; The robot is able to touch and move objects; Cronbach’s  $\alpha = .790$ ). Lastly, we assessed users’ evaluation of what and how well

the robot perceives and reacts towards the users using seven items which showed very good internal consistency (e.g., The robot was able to perceive my behavior; The robot was not able to understand my behavior; The robot reacted appropriately to my behavior; The robot did not perceive my behavior ; Cronbach's  $\alpha = .840$ ).

**Communicative Capabilities.** We assessed the robots perceived *Verbal Capabilities* with five ad-hoc generated items asking how well the robot understood and produced verbal contributions in the conversation (e.g., The robot understood me well; The robot did not hear me; The robot did not understand me ; Cronbach's  $\alpha = .840$ , cf. [15]). Moreover, we measured the robot's capabilities to understand and produce nonverbal behavior. *Nonverbal Understanding* was measured using six items (e.g., The robot noticed my gestures; The robot noticed my facial expressions; The robot was able to interpret my facial expressions and gestures correctly ; Cronbach's  $\alpha = .802$ ), and *Nonverbal Production* was measured using three items (The robot's gestures were expressive; The robot's gestures were unambiguous and clear to me; I understood the robot's gestures; Cronbach's  $\alpha = .849$ ).

#### 2.4 Dependent Variables: Linguistic Alignment

Participants played two structured games (guessing game and search game) in which the tutor and the participant took turns in constructing sentences. The verbal utterances of these games were analyzed regarding users' linguistic alignment with the artificial tutor. We used the exact same coding procedures as in our previous work (cf. [16]). Participants' utterances during the interaction with the robot were transcribed into plain text and coded along a predefined coding-scheme. All transcripts were coded by two individual coders and the inter-coder-reliability demonstrated a satisfying accordance between both coders (Guessing Game Cohen's Kappa = .948; Search Game Cohen's Kappa = .861).



**Fig. 2.** Left: Examples for interaction cards in the guessing game; Right: Example card for the search game

**Guessing Game.** The first structured game was a dialog based game (based on [4, 16]) in which participants took turns with the tutor in guessing two persons and their interaction on so-called interaction cards (cf. Fig. 2) by asking only yes-or-no questions (e.g. “Is the person on the left side female?”; “Is the person on the right side old?”, “Is the interaction between the two friendly?”). There were two rounds of guessing in which the system first guessed the participant’s card and then the participant guessed the system’s card. The system’s utterances between the two rounds varied in lexical choices when describing the features of the displayed characters (age (old vs. advanced in years), gender (male/female vs. a man/ a woman), facial hair (mustache vs. beard)). Moreover, the system used different verbs (has vs. wears), adjectives (friendly vs. kind) and syntactical constructions (person on the left side vs. the left person; active vs. passive). In total, we introduced seven variations. Participants’ verbal utterances were analyzed with regard to their lexical choices. A ratio was built for alignment for each of the seven aspects.

**Search Game.** In the second structured game participants and the tutor took turns in describing picture cards to one another by forming sentences based on the two characters (e.g. policeman and cowboy), the verb (e.g. to give), and the object (e.g. balloon, cf. Fig. 2) displayed on the card (cf. [4, 16]). Participants had two sets of cards (reading cards and search cards). The task was to take a card from the first card set (reading cards) and to form a sentence based on the characters, verb and object displayed on the card (e.g., The balloon was given to the policemen by the cowboy). The interaction partner’s task was to search in their set of “search cards” for this exact card. The system began the interaction and built a sentence. The participant had to find the card and put it away and in turn had to take a card from the “reading” set and form a sentence so that the tutor can find the card in its (imagined) search card pool and put it away. In total, the system read out 15 cards, thereby formed 15 sentences in three “blocks”. The first block i.e. the first five sentences were formed as passive voice, the second five sentences as prepositional phrase and the last as accusative. Three ratios were built for syntactical alignment within the three blocks.

## 2.5 Dependent Variables: Learning Outcome

Participant’s German language skills before and after the interaction were assessed with a so called C-Test ([www.c-test.de](http://www.c-test.de)), a cloze test which addresses language skills with regard to different dimensions. Scores were compared to test for learning gain after interaction. The C-Test has been used previously for assessing (improvement in) language skills. It usually comprises five short pieces of self-contained text (ca. 80 words each) in which single words are “damaged”. In order to reconstruct the sentences, participants have to activate their language fluency. Text pieces were taken from reading exams on an academic language level. Tests are analyzed by true-false answers. Participants could reach up to 100 points.

### 3 Results

#### 3.1 Evaluation of the Artificial Tutor

**Person Perception.** With two-factorial ANOVAs we tested the effect of the tutor’s embodiment and non-verbal behavior on users’ evaluations of the tutor’s Likability, Competence, and Autonomy (for descriptive data cf. Table 1). No main or interaction effects emerged for the three person perception scales.

**Table 1.** Means and standard deviations of the ratings of bodily-related capabilities and person perception of the tutor

	Bodily-related Capabilities			Person Perception		
	Perception	Mobility	Presence	Competence	Autonomy	Likability
	$M$ ( $SD$ )	$M$ ( $SD$ )	$M$ ( $SD$ )	$M$ ( $SD$ )	$M$ ( $SD$ )	$M$ ( $SD$ )
Virtual						
Idle	2.83 (.99)	1.22 (.39)	3.76 (1.19)	3.21 (.62)	2.21 (.95)	3.08 (.77)
nv	2.73 (.77)	1.64 (.69)	3.53 (.99)	3.52 (.74)	2.43 (.99)	3.26 (.89)
Robot						
Idle	3.02 (.84)	1.59 (.90)	4.19 (.77)	3.45 (.75)	2.17 (1.1)	3.13 (.94)
nv	3.07 (.97)	2.05 (.95)	4.16 (.86)	3.25 (.75)	2.22 (.98)	3.13 (.72)
Total	2.92 (.89)	1.63 (.81)	3.91 (.99)	3.36 (.72)	2.20 (1.0)	3.15 (.82)

**Bodily-Related Capabilities.** With two-factorial ANOVAs we tested the effect of the tutor’s embodiment and non-verbal behavior on users’ evaluations of bodily-related capabilities: (*Shared*) *Perception*, *Mobility*, and *Presence* (for descriptive data cf. Table 1).

The ANOVA on the dependent variable (*Shared*) *Perception* yielded no significant main effects and also no interaction effect for embodiment and non-verbal behavior. The results of the second ANOVA demonstrated significant main effects of embodiment ( $F(1,81) = 5.45$ ,  $p = .022$ ,  $\eta_p^2 = .06$ ) and nonverbal behavior ( $F(1,81) = 6.86$ ,  $p = .011$ ,  $\eta_p^2 = .08$ ) on the perceived *Mobility* of the tutor. Participants evaluated an embodied robot with higher mobility than a virtual agent. Moreover, a tutor using nonverbal behavior elicited a higher perception of mobility than a tutor that merely showed idle behavior. No interaction effect of embodiment and nonverbal behavior occurred for perceived mobility. For the dependent variable *Presence* there was a significant difference between the embodied robot and the virtual agent ( $F(1,81) = 6.30$ ,  $p = .014$ ,  $\eta_p^2 = .07$ ), while no main effect emerged for the robot’s behavior, and again, no interaction effect of embodiment and nonverbal behavior was present. Hence, participants who interacted with an embodied robot indicated higher presence of their interlocutor compared to those interacting with the virtual agent.



**Communicative Capabilities.** To investigate the effect of the experimental conditions on the perception of the tutor’s communicative capabilities, three two-factorial ANOVAs were calculated (for descriptive data cf. Table 2). For the dependent variable *Verbal understanding* a significant main effect emerged for the tutor’s embodiment ( $F(1,81) = 4.89, p = .030, \eta_p^2 = .06$ ). Participants indicated a stronger perception of verbal understanding for virtual agents than for robots. No main effect for the tutor’s nonverbal behavior occurred, nor did an interaction effect. There were also no main or interaction effects for the dependent variable *Nonverbal Understanding*. However, a main effect emerged for *Nonverbal Production*. The tutor displaying meaningful nonverbal behavior was rated higher with regard to Nonverbal Production than a tutor using idle behavior ( $F(1,81) = 8.30, p = .005, \eta_p^2 = .09$ ). The ANOVA revealed no significant effect of embodiment nor was an interaction effect found.

**Table 2.** Means and standard deviations of the ratings of communicative capabilities

	Communicative Abilities		
	Verbal understanding	Nonverbal understanding	Nonverbal production
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Virtual idel	4.43 (.52)	2.07 (.79)	2.49 (1.24)
Virtual nv	4.31 (.64)	2.29 (.88)	3.24 (1.20)
Robot idel	4.17 (.60)	2.49 (.76)	2.81 (1.23)
Robot nv	3.93 (.87)	2.50 (.1.0)	3.56 (1.11)
Total	3.01 (.69)	2.34 (.88)	3.03 (1.24)

### 3.2 Linguistic Alignment and Learning Outcome

**Guessing Game.** With the guessing game we examined participants’ syntactical and lexical alignment during the interaction. As described above, ratios were calculated for alignment (usage of the same lexical/syntactical choice (e.g. lexical choice *mustache*)/occurrence of the concept (e.g. number of expressions referring to a beard)). To examine whether embodiment or nonverbal behavior affects participants’ linguistic alignment, we conducted ANOVAs with both factors as independent variables and the seven ratios for linguistic alignment as dependent variables. There were no significant differences between the groups nor did we find significant interaction effects.

**Search Game.** The search game focused on the syntactical alignment. Regarding all 15 sentences, participants most often used accusative, followed by prepositional phrases and passive voice. In order to examine whether embodiment or nonverbal behavior of the artificial tutor affects participants’ syntactical alignment, we conducted ANOVAs with both factors as independent variables and

the alignment ratios. There were no significant differences between groups nor did we find significant interaction effects.

**Learning Outcome.** To explore whether the interaction has a positive effect on participants' language skills we analyzed the results of the C-Tests prior and after the interaction. Thus, we conducted split-plot ANOVAs with the group factors embodiment and nonverbal behavior and repeated measures for the C-Test scores. The scores did not differ between the two measuring points. Moreover, the experimental conditions showed no effect.

## 4 Discussion

With this work we contribute another puzzle piece to the ongoing debate of whether virtual agents or robots provide more benefits to the user. Previous work predominantly found that robots were superior over virtual agents (cf. [10, 14] for an overview). Still there is a lack of research on behavioral effects, particularly, with regard to linguistic behavior. This work presents an empirical study that investigates the effect of an artificial tutor's embodiment and nonverbal behavior on the tutor's perception, users' linguistic alignment and their learning outcome.

Our first central question to this study was whether an *artificial tutor's embodiment* (virtual agent vs. robot) influences participants' evaluation of the tutor, their lexical and syntactical alignment during interaction and their learning effect after the interaction (RQ1). Based on prior work, we hypothesized that the robot will be rated more positively than the virtual agent (H1). We did not find support for our hypothesis, since users' evaluations of the two systems did not differ with regard to perceived likability, autonomy or competence. However, we did find support for our assumption that evaluations regarding the perceived bodily-related and communicative capabilities of the tutor varies with embodiment (H2). Our results suggest that while a robot is perceived as being more present and having greater ability for mobility and tactile interaction than a virtual agent, the virtual agent was attributed higher communicative abilities with regard to verbal understanding. This is in line with the suggestions raised in the Embodiment and Corporeality Framework by Hoffmann et al. [10]. The Framework proposes that the core ability that is directly related to the physical embodiment of robots is Corporeality, i.e., the realism and material existence of the entity in the real world. Hence, the physical embodiment of the robotic artificial tutor should lead to higher ratings in corporeality and physical presence which is supported by the findings of the current study. The framework also proposes that evaluations of perceived bodily-related capabilities are additionally influenced by moderating variables such as the interaction scenario. Although communicative abilities such as verbal understanding, are not part of the EmCorp Framework, a similar proposal can be stated to explain the finding that the virtual agent is perceived as more verbally capable. The majority of the interaction with the participants was conversational, not involving the manipulation of objects. Although participants were using game material, the

task itself was not of physical nature like, for instance, in the Tower of Hanoi task in which objects have to be moved as the main task, but it was to produce sentences. Prior work has demonstrated that virtual agents seem to be preferred over physically embodied robots for purely conversational tasks (cf. [11]). Our finding that the virtual agent was attributed higher communicative abilities with regard to verbal understanding supports this assumption. With regard to participants' linguistic behavior and alignment tendencies, we found that the tutor's embodiment showed no effect. Users align toward virtual agents and robots in the same way. Moreover, both forms of embodiment elicited similar learning outcomes. This is against a body of research that predicts benefits of robots over virtual agents. At the same time, these results are in line with prior work that investigated the effect of embodiment on linguistic alignment in the setting of second language-acquisition [16] confirming this previous result that embodiment has no influence on users alignment and learning outcomes. In consequence, linguistic tutors do not need to be embodied as robots, which can save resources and make artificial tutors more accessible.

The second central research question was concerned with the effect of the artificial tutor's nonverbal behavior on users' linguistic behavior (RQ2). Again, no differences in participants' linguistic alignment and learning outcome was found indicating the tutor's nonverbal behavior did not influence how people align towards the tutor. Further, we assumed that the display of meaningful nonverbal behavior (in contrast to subtle idle behavior) leads to more positive evaluations (H3). Likewise to the non-existent effect of embodiment no positive effects were found for evaluations regarding likability, autonomy and competence. However, the nonverbal behavior of the tutor affected the perceived bodily-related and communicative capabilities. Those tutors using nonverbal behavior elicited a stronger perception of mobility and tactile interaction and a stronger production of nonverbal behavior. These effects are again in line with the EmCorp Framework [10] which suggests that the display or the absence of nonverbal behavior serves as a mediator for perceived capabilities, especially for agents who are technically able to show, for instance, gestures and in interaction settings in which the usage of gestures is helpful (e.g. pointing to game material).

The current study as well as our prior work [16] indicate on the one hand that the idea of exploiting human's tendencies of computer talk for SLA is not an effective measure, although participants in both studies showed great interest in using the robot or virtual agent for language training. On the other hand, both studies provided valuable insights into the effects of design decisions for artificial tutors such as embodiment and expressive nonverbal behavior.

## References

1. Atkinson, D., Churchill, E., Nishino, T., Okada, H.: Alignment and interaction in a sociocognitive approach to second language acquisition. *Mod. Lang. J.* **91**(2), 169–188 (2007). <https://doi.org/10.1111/j.1540-4781.2007.00539.x>

2. Bartneck, C.: Interacting with an embodied emotional character. In: Forlizzi, J., Hanington, B. (eds.) Proceedings of the 2003 international conference on Designing pleasurable products and interfaces, p. 55. ACM, New York, NY (2003). <https://doi.org/10.1145/782896.782911>
3. Bergmann, K., Branigan, H.P., Kopp, S.: Exploring the alignment space. Lexical and gestural alignment with real and virtual humans. *Front. ICT* 2 (2015). <https://doi.org/10.3389/fict.2015.00007>
4. Branigan, H.P., Pickering, M.J., Cleland, A.A.: Syntactic co-ordination in dialogue. *Cognition* 75(2), B13–B25 (2000). [https://doi.org/10.1016/S0010-0277\(99\)00081-5](https://doi.org/10.1016/S0010-0277(99)00081-5)
5. Branigan, H.P., Pickering, M.J., Pearson, J., McLean, J.F.: Linguistic alignment between people and computers. *J. Pragmatics* 42(9), 2355–2368 (2010). <https://doi.org/10.1016/j.pragma.2009.12.012>
6. Brennan, S.E., Clark, H.H.: Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Memory Cogn.* 22(6), 1482–1493 (1996). <https://doi.org/10.1037/0278-7393.22.6.1482>
7. Fasola, J., Mataric, M.: A socially assistive robot exercise coach for the elderly. *J. Hum.-Robot Interact.* 2(2) (2013). <https://doi.org/10.5898/JHRI.2.2.Fasola>
8. Fischer, K., Lohan, K.S., Foth, K.: Levels of embodiment. In: Yanco, H. (ed.) Proceedings of the seventh annual ACM/IEEE International Conference on Human-Robot Interaction, p. 463. ACM, New York, NY (2012). <https://doi.org/10.1145/2157689.2157839>
9. Giles, H.: Accent mobility: a model and some data. *Anthropol. Linguist.* 15, 87–105 (1973)
10. Hoffmann, L., Bock, N., Rosenthal von der Pütten, A.M.: The peculiarities of robot embodiment (EmCorp-Scale). In: Kanda, T., Abanović, S., Hoffman, G., Tapus, A. (eds.) HRI'18, 5–8 March 2018, Chicago, IL, USA, pp. 370–378. ACM, New York, NY, USA (2018). <https://doi.org/10.1145/3171221.3171242>
11. Hoffmann, L., Krämer, N.C.: Investigating the effects of physical and virtual embodiment in task-oriented and conversational contexts. *Int. J. Hum.-Comput. Stud.* 71(7–8), 763–774 (2013). <https://doi.org/10.1016/j.ijhcs.2013.04.007>
12. Kidd, C.D., Breazeal, C.: Effect of a robot on user perceptions. In: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3559–3564. IEEE, Piscataway, N.J (2004). <https://doi.org/10.1109/IROS.2004.1389967>
13. Kiesler, S., Powers, A., Fussell, S.R., Torrey, C.: Anthropomorphic interactions with a robot and robot-like agent. *Social Cogn.* 26(2), 169–181 (2008). <https://doi.org/10.1521/soco.2008.26.2.169>
14. Li, J.: The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum.-Comput. Stud.* 77, 23–37 (2015). <https://doi.org/10.1016/j.ijhcs.2015.01.001>
15. Rosenthal-von der Pütten, A.M., Krämer, N.C., Herrmann, J.: The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *Int. J. Soc. Robot.* 10(5), 569–582 (2018). <https://doi.org/10.1007/s12369-018-0466-7>
16. Rosenthal-von der Pütten, A.M., Straßmann, C., Krämer, N.C.: Robots or agents – neither helps you more or less during second language acquisition. In: Traum, D., Swartout, W., Khooshabeh, P., Kopp, S., Scherer, S., Leuski, A. (eds.) IVA 2016. LNCS (LNAI), vol. 10011, pp. 256–268. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-47665-0\\_23](https://doi.org/10.1007/978-3-319-47665-0_23)
17. Wunderlich, H.: Talking like a machine?! Linguistic alignment of native-speakers and non-native speakers in interaction with a virtual agent. Bachelor Thesis, University of Duisburg-Essen (2012)