Using Linguistic Alignment to Enhance Learning Experience with Pedagogical Agents: The Special Case of Dialect

Vivien Kühne, Astrid Marieke Rosenthal-von der Pütten, and Nicole C. Krämer

University of Duisburg-Essen, Department of Social Psychology:
Media and Communication, Forsthausweg 2, 47048 Duisburg, Germany
`vivien.kuehne@stud.uni-due.de`,
`{a.rosenthalvdpuetten,nicole.kraemer}@uni-due.de`

Abstract. Empirical research showed that verbal and nonverbal alignment occurs in HCI in the same way as in HHI [1-3]. Against the background of similarity attraction [4], a "we-feeling" within dialect-origin [5] and different investigations regarding speaking variations [6,7], the present study analyses the effect of the dialectical language usage of a virtual pedagogical agent within a tutoring setting and the ramifications for the learning situation. An experimental study with a between subject design (N=47) was conducted in which the virtual interlocutor explained and subsequently questioned the subjects about medical topics in either dialect or High German (via Wizard-of-Oz-scenario). The results show that linguistic alignment occurs in both conditions, but even more in interaction with the High German-speaking agent. Furthermore the dialect-using agent was rated as more likable while there were no effects with regard to social presence. Implications for theory and development are discussed.

Keywords: ECA, virtual agent, experimental study, linguistic alignment, communication adaptation theory, High German and dialect, social effects.

1 Introduction

In the 21th century we face virtual agents in online shops, airports, in museums and increasingly often in learning systems. Especially in the case of pedagogical agents, it seems to be important to use virtual characters which are as close to the learner as possible to increase learning motivation [8]. A possible characteristic which has not been considered so far is the use of dialect. Since dialect usage can establish a "we-feeling" [5] between people of the same dialect-origin and because people tend to like people similar to themselves (principle of similarity attraction; [4]), the usage of dialect in human-agent interactions could have the potential to bolster up motivational effects of pedagogical agents. According to Krämer and Bente [9], social factors have important influence on the interaction with a pedagogical agent in a tutoring situation.

Recent studies have shown that people align their nonverbal and verbal behavior in interactions with virtual agents just as they do in human-human interaction (HHI), e.g. regarding the amount of words used during conversation [3] or participants' smiling

behavior [2]. With regard to pedagogical agents, Rosenthal-von der Pütten et al. [10] investigated the linguistic alignment towards a tutoring agent by varying the lexical word-choice and found that participants align both to the agent using everyday language and the agent using technical terms. However, dialect has not been addressed so far. On that account, the present study includes the social factor of dialect and investigates linguistic alignment towards a dialect-speaking agent (Rhine-Ruhr region) and the differences between the ramifications of a dialect- or High German-agent. The possible alignment, the evaluation as well as a possible change of test-anxiety before and after the interaction may give further suggestions for the design and presentation of virtual agents and in this context, for designing better tutoring-systems.

1.1   Dialect in the Rhine-Ruhr Region

The standard language of any country serves as the main medium for communication of a culture, provides rules for syntax and lexis, and serves as benchmark for this language [11]. However, every culture has its own everyday modifications of their language, called colloquial language and dialect, which are more common in daily life. They are not only important because of the communicative function; they also indicate a membership to some kind of group (geographically and mentally) and therefore have a social function like creating a "we-feeling" [5]. This similarity attraction between people manifests not only in language, but also in attitude and values [4].

1.2   Linguistic Alignment in Human-Human Interaction

People use the appearance, nonverbal and verbal behavior of others to derive conclusions for an appropriate interaction and to find a shared communication concept [12]. According to Brennan [13], people have a "vocabulary problem", because they have thousands of possibilities to express themselves. An easy way to find a solution is alignment. People generate "conceptual pacts" [13] and use the strategy of convergence [14] or linguistic alignment [1], a dynamic and adaptive exchange between communication partners for a successful interaction, to manage the comprehension of an interaction. Branigan et al. [1] distinguish between conscious and unconscious alignment which both can occur in one interaction. For instance, people tend to repeat the things they have just said or heard and to align to single words (lexical alignment), sentences and structures (syntactical alignment), and to nonverbal features of speech like speech rate, pauses and pronunciation (for a review cf. [1]). Also of great importance is the linguistic alignment in the context of geographical and cultural conditions. With regard to certain dialects, people can align because of sympathy or integration, or they can refuse to align because of antipathy [6]. Till now it is underresearched how we can exploit this social aspect of language to enhance learning situations. In sum, linguistic alignment in HHI has different facets and includes variables which are not easily controllable. Human-computer interaction (HCI) and human-agent interaction (HAI) enable researchers to control these variables and facilitate systematic investigations which might lead to a better understanding of the processes and effects of linguistic alignment [15].

1.3   Linguistic Alignment in Human-Computer Interaction

That humans align with their interaction partner is also observable in HCI and HAI. In HCI research it was found that people aligned with regard to speaking pauses, sentence structures, speech rate, the usage of personal pronouns, choice of words and an emphasized articulation towards a computer-system (cf. [1] for a review). Branigan et al. [16] showed in a series of experiments that people tend to align even more to computers than to humans. Moreover, the expertise of the computer systems plays an important role, since it was found that participants adapted more strongly to computers that were presented as less capable than to computers that were presented as more capable. With regard to virtual agents, it has also been shown that humans align with their virtual interlocutor. People mimicked the agent's smiling behavior [2], aligned their amount of words during an interview [3], and showed lexical alignment [10]. The aspect of dialect has not been addressed in HAI so far, although some researchers examined speech inconsistencies and variations like interruptions, repetitions [17] or the usage of fillers like "uh" [18] and the alignment regarding these variations.

1.4   Research Questions

In the present study we explore whether participants linguistically align to a High German or dialect-speaking agent and how alignment affects the evaluation of the interaction. Based on previous findings on linguistic alignment in HAI we expect that in our study participants will also align to the virtual agent in their use of dialect or High German, respectively. We thus hypothesize that H1) participants, who are interacting with a dialect-speaking agent, use more dialectical words than participants who are interacting with a High German-speaking agent. Taking participants' natural tendency to use dialectal words into account measured by a previously assessed baseline, we hypothesize that H2a) participants in the dialect condition will use more dialectal words during than before the interaction and H2b) participants in the High German condition will use fewer dialectal words during than before the interaction.

Against the background of similarity attraction [4] and the "we-feeling" of "dialectal in-groups" described by Bichel [5] and the assumption that people are motivated to talk to others, who are similar to themselves, by appearance, origin or language, the following hypotheses are posited: H3a) The dialect-speaking agent will be rated more likable than the High German-speaking agent and H3b) participants will report higher social presence when interacting with the dialect-speaking agent compared to the High German agent. Moreover, we expect that H4) the dialect-speaking agent will be rated more positively by those participants who more frequently use dialect themselves compared to those who use dialect less frequently and that H5) the conversation with the dialect-speaking agent is rated more positively than the conversation with the High German-speaking agent.

Furthermore, we expect that our system might lead to a better tutoring and, regarding the coverstory mentioned below, is successful in reducing test anxiety. Thus we expect that H6) participants will report less test-anxiety after the interaction with the system than before.

2   Method

2.1   Experimental Design and Independent Variables

We used a one-factorial between subjects design with either a High German- or dialect-speaking agent. The utilized Rhine-Ruhr regional dialect is one of the most common in Northern Germany and is strongly influenced by the industrialization and the polish migration. Possible modifications are short-cuts of pronouns (e.g. „sie" = „se") or articles (e.g. "ein" = "en") as well as different pronunciations (e.g. "Alltag" = "Alltach"). In the present study, we varied the usage of those pronouns, articles and prepositions. For every sentence we altered two or three words from standard High German to dialect. We chose a female avatar (cf. Fig 1) by the Charamel Company (www.charamel.com), which was controlled by the investigator (Wizard-of-Oz-scenario). The nonverbal behavior of the agent (e.g. idle behavior like blinking and posture shifts, gestures) has been kept constant between conditions.

Fig. 1. The virtual avatar Gloria

2.2   Dependent Variables

The dependent variables used in this experiment were the *linguistic alignment of the participant, the perception of the agent, the social presence, the evaluation of the interaction in general* and *the situational test-anxiety of the participant.* These variables were measured by quantitative analysis (online questionnaires before and after the interaction, verbal behavior during the interaction).

Linguistic Alignment. Participant's verbal behavior before and during the interaction was recorded and transcribed. The used dialectical words were counted and dialect ratios were calculated (dialectical words/total of used words) for each participant for the interaction with the agent and for a previously surveyed baseline.

Perception of the Agent. The likability of the agent was measured using a semantic differential with 8 bi-polar pairs of adjectives (friendly-unfriendly, likable-unlikable,

pleasant-unpleasant, honest-dishonest, nice-mean, warmhearted-cold, compassionate-unconcerned, committed-uncommitted) which were rated on a 5-point Likert-scale (Cronbach's α = .844). In addition, we assessed as how pleasant the interaction with the tutor was evaluated. Participants rated on a 5-point Likert-scale from "displeasing" to "pleasing".

Social Presence. Social presence was assessed with two subscales from Nowak and Biocca's [19] social presence questionnaire: the subscale *perceived other's copresence* with 12 items (Cronbachs's α = .713) and the subscale *self-reported copresence* with six items (Cronbach's α = .578) were measured on a 5-point Likert-scale (from "agree" to "disagree"). The scale *self-reported copresence* was excluded from further analysis due to low internal consistency.

Evaluation of the Interaction. For the evaluation of the interaction with the virtual agent we used the *positive-evaluation-scale* with six items (e.g. "It was interesting to communicate with the tutor") rated on a 5-point Likert-scale (from "disagree" to "agree", Cronbach's α = .882).

Situational Test-anxiety. According to the cover story (see below), the participants were asked about their current feeling in this test-situation [20] before and after the interaction. To determine test-anxiety, a 7-point Likert-scale was used with eight adjectives (afraid, excited, uncertain, worried, tense, scary, fearful, and nervous; Cronbach's α = .960).

2.3 Participants and Procedure

A total of 47 persons (24 female) participated in the study with a mean age of 22.9 years (SD=1.78) ranging from 20 to 28 years. They were recruited via general advertising on campus and online advertising. The study was announced as an evaluation of new software that should help students to reduce test-anxiety in oral exams. To be included in the study participants had to fulfill the precondition of growing up in the Rhine-Ruhr-region. Upon arrival participants signed informed consent and received instructions. First, they filled in the pre-questionnaires (demographics, personality and situational test-anxiety) while the investigator ostensibly left the room in order not to disturb the participant during interaction. However, she retreated to another room from which she could control the interaction using pre-built scripts in a Wizard-of-Oz-scenario guaranteeing the same interactions in both conditions. After finishing the questionnaire participants put on a head-set and started the interaction on a second computer screen by saying "start". To generate a baseline for their natural usage of dialectal words, participants were first requested to talk three minutes about a random topic (e.g. the first week in university) in order to "adapt the computer to their speaking-characteristics". Subsequently, the interaction with the agent started which was divided in two sessions. In the first session, participants were told to play a word-understanding game with the agent in order "to train the agent to their voice". In this game they should move different objects (e.g. box, paper) according to the 15 instructions of the agent (e.g. "Open the box. Put the

stuffed animal into the box…"). After every move, participants should describe what exactly they were doing. During the second part of the interaction, the agent reported about the diseases Diabetes and Alzheimer (short version of [10]) and explained the clinical picture, the causes and effects of each disease. Directly after all of the 14 explanations, it asked questions about the recently specified content simulating a test situation. We used only open questions to avoid yes/no-answers. After the interaction, the participants had to fill in the second part of the questionnaire (situational test-anxiety, evaluation of the agent and interaction) after which they were fully debriefed and thanked for their participation.

3   Results

Linguistic Alignment. To examine whether participants who were interacting with a dialect-speaking agent used more dialectical words than participants interacting with a High German-speaking agent (H1), we conducted a one-way ANOVA with language usage of the agent as independent and dialect usage of the participant during the interaction as dependent variable. As expected, we found a main effect of the agent's language use (High German vs. dialect). Participants who interacted with the dialect-speaking agent used more dialectical words than participants who talked with the High German-speaking agent (cf. Table 1). To take into account the natural dialectical word usage of the participants, we conducted an analysis of covariance with the dialectical usage of the baseline as the covariate. There was also a significant effect of the agent's language usage on participants' usage of dialectal words after controlling for participants' usual dialectal usage ($F_{(2; 45)}= 18.969$; $p= <.001$; partial eta$^2= .463$). Furthermore, the differences within the groups before and during the interaction were investigated and main effects were found for both groups (H2a & H2b). The one-way ANOVAs revealed that participants who talked with the High German-speaking agent used fewer dialectical words during the interaction than before (cf. Table 2). Participants interacting with the dialect-speaking agent used more dialectical words during the interaction than before.

Perception of the Agent. To test whether the dialect-speaking agent was perceived as being more likable (H3), we conducted a one-way ANOVA with the language use of the agent as independent variable and the likability of the agent as dependent variable. As expected, the dialect-speaking agent was evaluated significantly more likeable than the High German-speaking agent (cf. Table 1). A one-way ANOVA was calculated with the dependent variable pleasantness of the interaction which revealed no significant effects. There was no correlation of the likability of the dialect-speaking agent and the own natural dialect usage (H4).

Social Presence. In order to examine whether the dialect-speaking agent elicits more social presence, a one-factorial ANOVA was calculated which showed no significant effect.

Perception of the Interaction. A one-way ANOVA with the *positive-evaluation*-scale as dependent variable and agent's language usage as independent variable

revealed a main effect contrary to H5: The interaction was evaluated more positive when talking to the High German- than to the dialect-speaking agent (cf. Table 1).

Situational Test-anxiety. The test-anxiety mean values indicate a lower situational test-anxiety after the interaction with the system (Before: *M*=3.16, *SD*=1.634; After: *M*=2.88, *SD*=1.577). However, the one-way ANOVA with repeated measurement for the factor situational test-anxiety showed no significant effect.

Table 1. ANOVA - Means and standard deviations for the used dialectical words during the interaction, the likability of the agent, and the positive evaluation of the interaction

	High German	Dialect				
	M (SD)	*M (SD)*	F	p	ηp²	df
Used dialectical words in game	4.97 (3.97)	9.91 (6.26)	10.570	.002	.190	1
Used dialectical words in test	7.58 (2.11)	10.49 (2.05)	22.895	<.001	.337	1
Used dialectical words total	6.39 (1.93)	10.31 (3.32)	24.841	<.001	.356	1
Likability	2.35 (.72)	2.73 (.53)	4.058	.050	.083	1
Positive evaluation of interaction	3.37 (.84)	2.78 (.86)	5.220	.027	.104	1

Table 2. ANOVA - Means and standard deviations for the used dialectical words before and during the interaction

Condition	Baseline	Interaction				
	M (SD)	*M (SD)*	F	p	ηp²	df
High German agent	11.23 (3.082)	6.39 (1.931)	59.369	.000	.721	1
Dialect agent	8.43 (3.237)	10.32 (3.318)	7.334	.013	.250	1

4   Discussion

We presented an experimental study examining the effects of dialect in HAI regarding linguistic alignment, the likeability of the agent and the perception of the interaction. Linguistic alignment was found in both conditions. People who talked with the High German-speaking agent used fewer dialectical words, those who talked to the dialect-speaking agent used more dialectical words than usually. Our results support existing findings which have shown linguistic alignment in HCI for instance by Branigan et al. [1] and Brennan [13]. Moreover, the results are in line with the findings of Ferguson [21] and Giles [6] indicating that alignment in pronunciation and accent is used to guarantee a better understanding. Comparing both conditions, people who talked with the High German-speaking agent linguistically aligned to a greater extent than people who talked to the dialect-speaking agent. This is a quite intuitive result, because in general dialect-speaking Germans are used to adapt to High German when speaking with Germans not from their dialectal region. An explanation for the distinct

alignment in the High German condition might be the possible weak manipulation of the dialect-speaking agent, who spoke a rather light version of the Rhine-Ruhr dialect. A stronger dialect might have revealed stronger alignment in the dialect condition, but this is uncommon in an academic setting and rather used in rather socially disadvantaged classes. However, all effects were persistent when controlling for participants' usual dialect usage (assessed by a baseline). Besides the actual alignment in participants' behavior, we assumed that dialect adds a social component to the interaction resulting in positive social effects. Indeed, the dialect-speaking agent was rated more likable than the High German-speaking agent. This supports the assumption that using dialect creates a "we-feeling" [5] between the agent and the participants. In accordance with the similarity-attraction theory [4] this shows a higher positive attitude towards the in-group agent. Although the interaction with both agents was rated as rather pleasant, dialectical usage seems to have no significant effect concerning the perception of the pleasantness of the interaction. Furthermore it seems that the mentioned similarity attraction and the "we-feeling" also have no significant effect towards the perception of social presence. In this context, another explanation could be a perceived discrepancy between behavior and appearance. Perhaps, the serious appearance of the agent didn't fit to the rather chummy verbal behavior as which the Rhine-Ruhr dialect might have been perceived. Consistency across modalities and other factors (like behavior and appearance), however, is an important factor for successful communication [22,23]. Surprisingly, the interaction itself was evaluated more positively when interacting with the High German-speaking agent. This might be due to the test situation. In academia, people are used to be in test situations with High German conditions. Thus, it might be confusing to be questioned by a dialect-speaking interlocutor, who, moreover, is dressed in a business suit. With regard to the assessed test-anxiety, no significant decrease in test-anxiety after the interaction was found, although the descriptive data indicate a reduction of test-anxiety. A considerable reduce of test-anxiety is presumably only possible when participants use the system more frequently [10]. Moreover, the test situation was interweaved with tutoring sessions where the agent presented information (in contrast to real oral exams) and participants are aware that the situation is fictitious and has no real consequences.

Limitations and Perspectives. One limitation of this study is that participants showed a significantly different dialect usage between the two groups before the interaction. This limitation could be avoided, for instance, by additional containment of the examined dialectal region or questions about the intensity of dialect usage. Another possibility is to divide the experiment into two sessions and assign participants in the second experimental session based on results of the previously assessed baseline. Additionally, it has to be mentioned that the structure of the conversation between participant and agent was highly structured to some extent. That is why it could be interesting to analyze free speech of the participants. Regarding the assumption of Branigan et al. [1] that alignment occurs even more in HCI than in HHI, it might further be interesting to compare the dialectical alignment with a HHI condition to draw conclusions on the different degrees of alignment in HCI and HHI. Moreover, different dialects or lighter and stronger versions of the

same dialect could be used to determine the effects of this social language component and examine how virtual tutoring systems can benefit from using dialect. In this context the durability and sustainability of alignment and the effects on evaluation over a longer period of system usage are further interesting fields of examination. This long-term evaluation would also be interesting for the effects on reducing test-anxiety. With respect to the tutoring system itself, some suggestions for improvement can be proposed. The described context of the learning situation and the agent's appearance should be taken into account. While it seems to be appropriate for a virtual examiner to speak High-German and be dressed rather formally, a tutoring agent which solely supports learning might benefit from a more informal appearance and the use of dialect. Another limitation was the rather restricted mimic feedback of the agent, which is meaningful for social interaction with an agent [24]. Moreover the effect of linguistic alignment (dialectical and standard) on knowledge transfer should be tested in future work in order to design an effective tutoring system.

References

1. Branigan, H.P., Pickering, M.J., Pearson, J., et al.: Linguistic alignment between people and computers. Journal of Pragmatics 42(9), 2355–2368 (2010), doi:10.1016/j.pragma.2009.12.012
2. Krämer, N.C., Kopp, S., Becker-Asano, C., et al.: Smile and the world will smile with you - The effects of a virtual agent's smile on users' evaluation and behavior. International Journal of Human-Computer Studies 71(3), 335–349 (2013), doi:10.1016/j.ijhcs.2012.09.006
3. von der Pütten, A.M., Hoffmann, L., Klatt, J., Krämer, N.C.: Quid Pro Quo? Reciprocal Self-disclosure and Communicative Accomodation Towards a Virtual Interviewer. In: Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R. (eds.) IVA 2011. LNCS, vol. 6895, pp. 183–194. Springer, Heidelberg (2011)
4. Aronson, E., Wilson, T.D., Akert, R.M.: Social psychology, 7th edn. Prentice Hall, Upper Saddle River (2010)
5. Bichel, U.: Problem und Begriff der Umgangssprache in der germanistischen Forschung. Max Niemeyer Verlag, Tübingen (1973)
6. Giles, H.: Accent mobility: A model and some data. Anthropological Linguistics 15, 87–105 (1973)
7. Iacobelli, F., Cassell, J.: Ethnic Identity and Engagement in Embodied Conversational Agents. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D., et al. (eds.) IVA 2007. LNCS (LNAI), vol. 4722, pp. 57–63. Springer, Heidelberg (2007)
8. Moreno, R., Flowerday, T.: Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. Contemporary Educational Psychology 31(2), 186–207 (2006), doi:10.1016/j.cedpsych.2005.05.002
9. Krämer, N.C., Bente, G.: Personalizing e-Learning. The Social Effects of Pedagogical Agents. Educ. Psychol. Rev. 22(1), 71–87 (2010), doi:10.1007/s10648-010-9123-x
10. Rosenthal-von der Pütten, A.M., Wiering, L., Krämer, N.C.: Great minds think alike. Experimental study on lexical alignment in human-agent interaction. i-com (in press)
11. Fekeler-Lepszy, E.: Gesprochene Sprache im Ruhrgebiet. K. Farin & H.-J. Zwingmann (1983)
12. Garrod, S., Anderson, A.: Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. Cognition 27(2), 181–218 (1987), doi:10.1016/0010-0277(87)90018-7

13. Brennan, S.E.: Lexical entrainment in spontaneous dialog. In: Proceedings of the 1996 International Symposium on Spoken Dialogue, ISSD 1996, pp. 41–44 (1996)

14. Giles, H., Coupland, N., Coupland, J.: Accommodation theory: communication, context, and consequence. In: Giles, H., Coupland, J., Coupland, N. (eds.) Contexts of Accommodation: Developments in Applied Sociolinguistics, pp. 1–68. Cambridge University Press, Cambridge (1991)

15. Kuhlen, A.K., Brennan, S.E.: Language in dialogue: when confederates might be hazardous to your data. Psychonomic Bulletin & Review 20(1), 54–72 (2013)

16. Branigan, H.P., Pickering, M.J., Pearson, J., et al.: The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. Cognition 121(1), 41–57 (2011), doi:10.1016/j.cognition.2011.05.011

17. Brennan, S.E., Schober, M.F.: How Listeners Compensate for Disfluencies in Spontaneous Speech. Journal of Memory and Language 44(2), 274–296 (2001), doi:10.1006/jmla.2000.2753

18. Pfeifer, L.M., Bickmore, T.: Should Agents Speak Like, um, Humans? The Use of Conversational Fillers by Virtual Agents. In: Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjálmsson, H.H., et al. (eds.) IVA 2009. LNCS, vol. 5773, pp. 460–466. Springer, Heidelberg (2009)

19. Nowak, K.L., Biocca, F.: The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. PRESENCE: Teleoperators and Virtual Environments 12(5), 481–494 (2003), doi:10.1162/105474603322761289

20. Jacobs, B.: Fragebogen zur aktuellen Prüfungsangst: 288 (2010)

21. Ferguson, C.: Toward a characterization of English foreigner talk. Anthropological Linguistics 17(1), 1–14 (1975)

22. Nass, C., Gong, L.: Maximized modality or constrained consistency. In: Proceedings of the AVSP 1999 Conference, pp. 1–5 (1999)

23. Goetz, J., Kiesler, S.B., Powers, A.: Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation. In: Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2003), pp. 55–60 (2003), doi:10.1109/ROMAN.2003.1251796

24. Nass, C., Isbister, K., Lee, E., et al.: Truth Is Beauty: Researching Embodied Conversational Agents. In: Cassell, J. (ed.) Embodied Conversational Agents, pp. 374–402. MIT Press, Cambridge (2000)