The Effects of Culturally Congruent Educational Technologies on Student Achievement

Samantha Finkelstein[1], Evelyn Yarzebinski[1], Callie Vaughn[2], Amy Ogan[1], and Justine Cassell[1]

[1] Human-Computer Interaction Institute
[2] Language Technologies Institute Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh PA, 15289
{slfink,eey2,cllvaughn,aeo,justine}@cs.cmu.edu

Abstract. Dialectal differences are one explanation for the systematically reduced test scores of children of color compared to their Euro-American peers. In this work, we explore the relationship between academic performance and dialect differences exhibited in a learning environment by assessing 3rd grade students' science performance after interacting with a "distant peer" technology that employed one of three dialect use patterns. We found that our participants, all native speakers of African American Vernacular English (AAVE), demonstrated the strongest science performance when the technology used AAVE features consistently throughout the interaction. These results call for a re-examination of the cultural assumptions underlying the design of educational technologies, with a specific emphasis on the way in which we present information to culturally-underrepresented groups.

Keywords: culture, dialect, peer models.

1 Introduction

Despite the typically standardized nature of mainstream school experiences, children begin their educational journey with unique cultural backgrounds that impact how they speak, collaborate with their peers, interact with authority figures, and talk about school-relevant topics such as science [1; 2]. Indeed, students may encounter cultural and language mismatches with their teachers as early as pre-school [3], with teachers mistaking cultural difference as deficits, unwittingly perpetuating an academic achievement gap [4].

Increasingly, the persistently lower test scores of students of color as compared to their Euro-American peers have been attributed in part to dialectal differences between students [4; 5; 6]. For example, some (but not all) African American students may come to school speaking a stigmatized, non-standard dialect of English referred to as African American Vernacular English (AAVE) [7], which has a unique phonology, morphology, and syntax that is regularized across users [8; 9]. Though the exact mechanisms behind the phenomenon are unclear, students who come to school speaking this dialect consistently score lower on indices of emergent literacy skills

K. Yacef et al. (Eds.): AIED 2013, LNAI 7926, pp. 493–502, 2013.
© Springer-Verlag Berlin Heidelberg 2013

than their predominantly Mainstream American English (MAE)-speaking peers [10; 11; 12]. Researchers and teachers alike are unsure of how to address these sensitive issues in a classroom, and whether to insist students transition to a mainstream dialect, teach in the students' native dialect, or provide instruction in code-switching (switching between dialects in different contexts) [13]. Unfortunately, insufficient evidence currently exists to fully understand how these different language ideologies might affect the learning and well-being of students who speak with non-standard dialects – a necessary step in supporting them academically, while not denying them access to key parts of their identity [14; 15].

We believe that educational technologies that employ culturally-congruent designs [16] can not only provide insight about culture's role in learning, but also significantly reduce the achievement gap. Previous research documents the importance of language similarities in learning, with students learning best from teachers who have similar accents to their own [17] or when allowed to work on material with a partner in their native language [18]. The majority of previous culturally-sensitive educational technologies, however, have exclusively focused on modeling surface level traits such as skin color, ignoring deeper cultural phenomena associated with communication [19]. There is therefore a need for experimental manipulations of language practices within educational technologies to examine the effect of dialect congruence between the student and technology. As such, in this work, we address this substantial lacuna with what we belief to be the first comparison of student learning in the context of technology that speaks one of the three dialectal patterns discussed above: exclusively Mainstream American English (MAE), exclusively AAVE, or code-switching.

2 Related Work

A limited number of educational technologies have addressed the discontinuity between students' home culture and their school environment by integrating commonly perceived aspects of minority culture, such as rap songs or cornrow hairstyles, into educational software [20;21;22]. For example, Culturally Situated Design Tools teaches transformation geometry with plaited symbols that can be rotated to re-create examples of African American cornrow hairstyles [20]. Gilbert et al. [21] similarly developed AADMLSS, in which students watch an embodied virtual agent solve a series of math problems grounded in neighborhood tasks, with mathematical explanations provided through rap songs. These ideas are also employed in Lyric Reader [22] which uses child-appropriate rap to promote literacy. Despite the positive qualitative results of these technologies, most have been compared against a "worksheet" control, rather than a similar technology that exclusively manipulates the presence of the intended cultural stimuli, such as corn rows or rap lyrics.

Also noteworthy is research on cultural sensitivity with virtual agents, such as Hayes-Roth's description of how agents from different cultural backgrounds could use language to embody deep-seated cultural differences [23]. There have been some studies which have included dialect as one index of culture, although it was not manipulated as distinct from skin color, and no information about the frequency or features of the non-standard dialects were discussed [24]. More commonly, studies investigating the impact of cultural differences in agents neglect to manipulate dialect at all, such as Baylor et al. [25], who found that varying agent age, gender, and ethnicity (including African American) affected both student perceptions of the agent's intelligence, and their learning. However, the authors did not manipulate dialect, nor did they report whether AAVE was used for the voice of the African American agents.

In our previous work, we addressed some of these issues by examining performance in a collaborative bridge-building task where students were either partnered with a human classmate, or a virtual peer who code-switched between speaking AAVE during science collaboration and MAE during a presentation to the teacher task [26]. While most students reduced their use of AAVE during the presentation task, those who were partnered with a code-switching agent demonstrated a significantly greater reduction of AAVE during formal presentation. However, this earlier work only examined one particular dialect switching pattern (AAVE for collaboration, MAE for presentation), motivating our current work to experimentally compare the effects of three dialect switching patterns in an agent, patterns whose benefits are currently being debated [27].

3 Methodology

We worked with 29 3rd grade students at a low SES (99% free or reduced lunch) 100% African American urban charter school to address whether students who speak with a non-standard dialect would demonstrate greater science proficiency after interacting with an educational technology that used the same dialect features in its own speech. We eliminated six students from the analysis due to data loss. Classroom observations determined that all students spoke AAVE to varying degrees, and dialect use was sometimes openly called out and stigmatized by the teacher.

We designed what we call a Distant Peer paradigm, in which children were partnered with an agent throughout the study to make audio recordings of a social task (an introduction about the student's interests) and a science task (providing scientific hypotheses about a pair of fictional creatures). Children believed their agent partner attended "a local school just like [theirs]," had completed the task a few days earlier, and would be later receiving the recordings the children created (like a pen pal). The agent partner was represented by a gender-ambiguous African American character ("Jamie") shown on individual laptops (see Figure 1). Jamie's voice was pre-recorded by a confederate who was bidialectal in AAVE and MAE, with recordings pitch-shifted to sound like a child. Children were randomly assigned to condition: (1) MAE, with an agent partner who spoke in MAE during both the social and science tasks; (2) AAVE, a partner who used AAVE in both tasks; and (3) code-switching, a partner who code-switched from AAVE in the social task to MAE in the science task. We emphasize that the only difference between the AAVE and code-switching agents is the dialect in which children heard the agent's initial four minute social introduction, allowing us to examine if science performance would be affected by the agent's dialect even in previous social dialogue unrelated to the task.

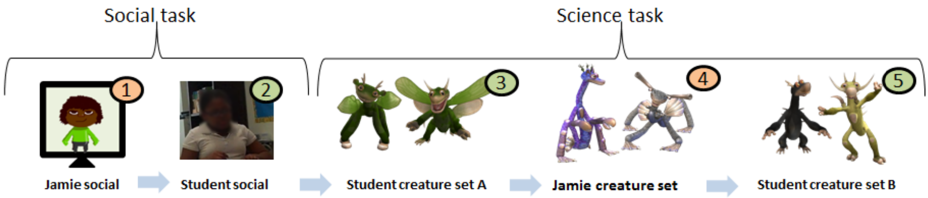

Fig. 1. Procedure: (1) listen to agent's social recording; (2) produce a social recording; (3) produce a first science recording; (4) listen to agent's science recording; (5) produce a second science recording. Order of creature sets A and B were counter-balanced.

In the science task, students were given pictures of fictional species in identical eco-systems. They were asked to "record [their] best hypotheses" about how the creatures might behave and interact within their environment for four minutes, both before and after hearing the peer science model, as shown in Figure 1. The open-ended nature of this task allowed students to monologue freely, allowing us to observe the students' use of dialect features and assess their science talk.

Jamie's social and science monologues were identical in both content and prosody across all condition [see previous work, 28], and differed only in presence of AAVE dialectal features (e.g. MAE: "the creatures don't have any claws" vs. AAVE: "the creatures don't have no claws.") For ease of exposition, specifics of Jamie's AAVE and science talk are described further in section 4.

4 Data Annotation

We focus our analysis on students' science talk and dialect use during the two four-minute science recordings students created (before and after hearing the agent model). The data was annotated by coders who were blind to condition. They achieved Cohen's Kappa agreement ratings of over .7 for each feature described below.

Our science annotation scheme was based on Linn et al.'s categorization of contribution, support, and complexity in science reasoning [29], as well as McNeill's description of claims and appropriate reasoning in science explanations [30]. Our science coding manual was reviewed and iterated upon with our science teacher advisor to obtain construct validity.

We first segmented students' monologues into units, defined as individual contributions that captured cohesive components of students' scientific ideas, as described in [31]. Each contribution was then coded for the presence of the following non-mutually-exclusive features: (1) claims, (2) reasoning, and (3) scientific integration (defined as integration of scientific ideas based in prior knowledge, analogies within ecology, or inferences about functionality.) Contributions that included at least one of each of these features (e.g. "the first creature is probably a carnivore because it looks fast and has sharp teeth and can use them to attack other animals for food") we called

"Strong Scientifically-Reasoned Arguments" (SSRAs) based on prior literature about elementary school level science arguments [29; 30]. Coders' inter-rater reliability for SSRAs was (κ = .92).

AAVE features were coded using the scheme proposed and validated by Renn [32], with slight modifications. We coded for morphosyntactic features, including multiple negation, copula deletion, and zero plural (see [37] for full list), as well as one phonetic feature, nasal fronting, identified as particularly relevant in children's code-switching [38]. While Renn additionally proposed two other phonetic features characteristic of AAVE, we primarily focus our analyses on morphosyntax because this has been shown to be more under children's control than their phonology, and therefore more likely able to be dampened when children code-switch [38]. We operationalize amount of dialect use with the Density Dialect Measure (DDM), calculated by dividing the total number of coded AAVE features used over the total number of words and multiplying by 100 as in [7].

Jamie's monologues in the AAVE condition included a subset of the 27 morphosyntactic features present in [32], because it would not have been realistic to fit examples of each feature into such small speech samples. The speech samples did contain a number of phonetic AAVE features because they were recorded by a natural bidialectal speaker, but we did not code for all of these features in our participants because of the difficulty of successfully annotating difficult phonetic features such as vowel quality. Jamie's monologues in the AAVE condition averaged a DDM of 13.3 and was designed to be substantially higher than our participants' (M = 1.5), such that there would be a clear distinction between MAE and AAVE conditions.

Jamie's science monologue included six examples of SSRAs, alongside other scientifically-relevant content, such as observations ("it looks like the creature has gills"), comparisons ("one creature looks like it can stand up on both legs, but the other one looks like it can only swim"), and questions "I wonder which one is more dangerous…").

5 Results

We operationalize students' science talk strength as the number of Strong Scientifically Reasoned Arguments (SSRAs) students provided in each four minute science recording. Jamie provided six examples of SSRAs (as well as other kinds of age-appropriate talk such as observations and comparisons of creatures) in the agent's 4 minute-long monologue. We first performed paired-samples t-tests to determine whether listening to Jamie's science talk recordings increased students' likelihood of producing on-task science contributions, SSRAs, reasoning, and scientific integration (ecological analogies, functionality, and prior knowledge) between their first and second science recordings, regardless of condition. As shown in Table 1, across all students the number of on-task science contributions, the number of SSRAs, and the amount of reasoning significantly increased from the first to second science recording. The incorporation of scientific integration did not change.

Table 1. Comparison of Students' Science Talk in First and Second Monologue

	Science 1 *M* (*SD*)	Science 2 *M* (*SD*)	*t*	*df*
# Contributions	15.35 (5.80)	18.65 (7.46)	-2.67*	22
# SSRAs	0.30 (.77)	1.83 (1.52)	-4.22***	22
# Reasoning	1.43 (2.62)	4.09 (3.07)	-4.46***	22
# Scientific Integration	2.91 (2.41)	3.96 (3.28)	-1.52	22

In order to test the hypothesis that students' ability to produce SSRAs would improve differentially based on condition, we ran a Repeated Measures ANOVA comparing the count of SSRAs in the first and second recording, with a between-subjects factor of condition. Results showed a significant main effect of science recording, $F(1, 20) = 26.06$, $p < .001$, showing that, as above, students increased their production of SSRAs after hearing a model. In addition, a significant interaction between condition and recording ($F(2, 20) = 6.887$, $p < .01$), revealed with Bonferroni post-hoc analyses that students in the AAVE condition showed a significantly higher increase than the MAE condition in production SSRAs from time one to time two ($p < .05$). The code-switching condition was not significantly different from either the AAVE or MAE condition at $\alpha = .05$, with gains between the other two conditions.

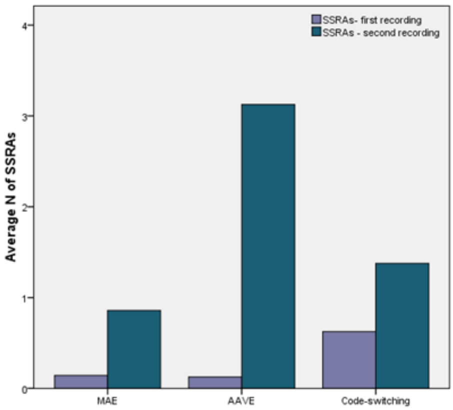

Fig. 2. Left: Relationship between students' initial DDM during the first science task and their subsequent performance on the second science task. **Right**: SSRAs produced by condition before and after interacting with Jamie.

A Repeated Measures ANOVA compared students' DDMs in the first (m – 1.5, range = 0 to 3.11) and second (m = 1.3, range = 0 – 4.5) science recording, with a between-subjects factor of condition. We clarify that demonstrating a DDM of 0 in these particular tasks does not mean that these students are not speakers of AAVE, as students may use the dialect in different contexts. There was no significant DDM difference between students' first or second recording, with no effect of condition. While not significant, students in the MAE condition trended to reduce their AAVE (M_{Δ} = -.0039), students in the code-switching condition trended to increase their AAVE (M_{Δ} = .0024), and students in the AAVE condition trended to stay the same (M_{Δ} = .0002). We reiterate that Jamie's DDM at 13.3 was substantially higher than our participants'.

6 Discussion

Though the vast majority of technologies are designed to communicate information to students using a mainstream dialect, the results of this work demonstrate that the strongest improvements in science talk were seen among students who heard the technology speak in AAVE – the children's native dialect. We additionally found that students' own dialect patterns did not change from their first science recording to their second. This has important implications, as teachers worry that allowing the vernacular dialect into their classroom will perpetuate the consistent use of the vernacular among students, and make them even less likely to use the standard [27]. However, our study did involve children only hearing very limited samples of the agent's speech in monologue, and we may see stronger effects on students' dialect use over greater periods of time spent interacting with the system, or during continuous dialogues with the system. Furthermore, we note that code-switching is a very complex linguistic process, and that the dialectal model we provided was a simplified instantiation of this process. Future analyses will continue to iterate our language model to better represent the intricacies of fluid switching behaviors seen among actual bidialectal students.

Because of the complex relationship between dialect and education, we propose three potential explanations for our result that AAVE-speaking students demonstrate increased success with AAVE speaking technology. The first is that there is a reduction of cognitive load when working with systems that communicate in students' native dialect, as supported by previous research that demonstrates students learn best from teachers who share their accent [17]. Students fluent in the mainstream dialect may be able to expend less effort during a learning task translating the provided information into a format they can better understand. It may also be that students are better able to demonstrate learning if they feel comfortable producing it in their native dialect, as they may be after hearing an example of the information provided in such dialect. The second explanation could be that students felt an increased rapport, or sameness, with the agent in our system who spoke in their own dialect, as students typically learn from those who are more similar to themselves [33]. Our previous work examined the acoustic features of students' recordings by condition, and found that those with an AAVE-speaking agent spoke more loudly, more quickly, and with more pitch fluctuation during the social introduction task compared to their peers who had an MAE-speaking agent. This leads us to believe that students felt more comfortable with an AAVE-speaking partner, which may have facilitated learning. The final explanation is that students may have been attending more closely to a technology who spoke in AAVE due to a novelty effect, as they have likely never experienced a system to communicate in this dialect before. Future studies which analyze the use of this technology over time will provide more insight about how these potential explanations affect students' overall learning, and clarify the role that each plays in the students' educational process.

7 Conclusion and Future Work

In this work, we provide, to our knowledge, the first example of an educational technology that experimentally manipulates different dialectal patterns and investigates subsequent academic performance. We exposed AAVE-speaking 3rd graders to an educational technology that used one of three dialect switching patterns, and conclude with two primary results: (1) students demonstrate improvement in science talk after listening to a science model from a peer educational technology, and (2) improvement is greatest among AAVE-speaking children with a peer that speaks in AAVE.

Our future work will incorporate our results into our virtual peer technology [26], and investigate more complex models of dialect switching, as this is a complicated and socially-driven phenomenon. Within these evaluations we will additionally examine transfer, retention, and longitudinal effects of learning with culturally sensitive technologies, as well as the long-term social benefits of culturally similar peer technologies, such as improved self-efficacy.

We believe the results of this work provide two primary lessons. The first is that we can design technologies to provide insight into complex and sensitive phenomena which are not yet fully understood. The second is that we make culturally-charged decisions in the design of every aspect of our technologies, and these may have significant impacts on users from underrepresented populations. As it is unreasonable to expect young children to be able to accurately articulate how sensitive topics such as race, identity, and cultural affiliation in educational environments may affect their learning, developers can work towards culturally sensitive technologies by experimentally manipulating aspects of our work, and monitoring the effects on children. This process not only provides insight about how to best design technologies for our target audiences to promote educational and socio-emotional success, it also acts to serve as the ground on which we begin to identify what (and how) cultural factors play into students' experiences. This study demonstrates the critical effects of small decisions within a system, and calls for developers to question the assumptions they put forth in the development of their own systems.

Acknowledgements. This work was supported in part by The Heinz Foundation (# C2902 and # E0510), the Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education (# R305B090023), and the National Science Foundation (# 0946825). Many additional thanks to the RAs of the ArticuLab who've helped with this project!

References

1. Gutiérrez, K.D., Rogoff, B.: Cultural ways of learning: Individual traits or repertoires of practice. Educational Researcher 32(5), 19–25 (2003)
2. Lee, C.D.: Bridging home and school literacies: Models for culturally responsive teaching, a case for African American English. In: Handbook of Research on Teaching Literacy Through the Communicative and Visual Arts, pp. 334–345 (1997)
3. Michaels, S.A.: Sharing time: children's narrative styles and differential access to literacy. Doctoral dissertation, University of California, Berkeley (1991)
4. Atkinson, J.L.: Are We Creating the Achievement Gap? Examining How Deficit Mentalities Influence Indigenous Science Curriculum Choices. In: Cultural Studies and Environmentalism, pp. 439–446 (2010)
5. Morrison, F.J., Bachman, H.J., Connor, C.M.: Improving literacy in America: Guidelines from research. Yale University Press (2005)
6. Gann, R.R.: Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In: Handbook of Erly Literacy Research, pp. 97–110 (2001)
7. Craig, H.K., Thompson, C.A., Washington, J.A., Potter, S.L.: Phonological Features of Child African American English. Journal of Speech, Language and Hearing Research 55, 623–635 (2003)
8. Green, L.J.: African American English: a linguistic introduction. Cambridge University Press (2002)
9. Rickford, J.R., Labov, W.: African American vernacular English: Features, evolution, educational implications, p. 157. Blackwell, Oxford (1999)
10. Connor, C., Craig, H.: African American preschoolers' language, emergent literacy skills, and use of African American English: A complex relation. Journal of Speech, Language, and Hearing Research 49, 771–792 (2006)
11. Charity, A.H., Scarborough, H.S., Griffin, D.M.: Familiarity with school English in African American children and its relation to early reading achievement. Child Development 75(5), 1340–1356 (2004)
12. Scarborough, H.S.: Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In: Neuman, S.B., Dickinson, D.K. (eds.) Handbook of Early Literacy Research, pp. 97–110. Guilford, New York (2001)
13. Ball, A.F., Farr, M.: Language varieties, culture, and teaching the English. In: Handbook of Research on Teaching the English Language Arts, pp. 435–445. Lawrence Erlbaum, Mahwah (2003)
14. Terry, N.P., Connor, C.M., Thomas-Tate, S., Love, M.: Examining relationships among dialect variation, literacy skills, and school context in first grade. Journal of Speech, Language and Hearing Research 53(1), 126 (2010)
15. Connor, C.M.: Language and literacy connections for children who are African American. Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations 15, 43–53 (2008)
16. Mohatt, G., Erickson, F.: Cultural differences in teaching styles in an Odawa school: A sociolinguistic approach. Culture and the Bilingual Classroom: Studies in Classroom Ethnography 105 (1981)
17. Gill, M.M.: Accent and stereotypes: Their effect on perceptions of teachers and lecture comprehension (1994)
18. Webb, L., Webb, P.: Introducing discussion into multilingual mathematics classrooms: An issue of code switching? Pythagoras (67), 26–32 (2011)

19. Cassell, J.: Social practice: Becoming enculturated in human-computer interaction. In: Stephanidis, C. (ed.) UAHCI 2009, Part III. LNCS, vol. 5616, pp. 303–313. Springer, Heidelberg (2009)
20. Eglash, R., Bennett, A., O'Donnell, C., Jennings, S., Cintorino, M.: Culturally situated design tools: Ethnocomputing from field site to classroom. American Anthropologist 108(2), 347–362 (2006)
21. Gilbert, J.E., Arbuthnot, K., Hood, S., Grant, M.M., West, M.L., McMillian, Y., Eugene, W.: Teaching algebra using culturally relevant virtual instructors. The International Journal of Virtual Reality 7(1), 21–30 (2008)
22. Pinkard, N.: Rappin'Reader and Say Say Oh Playmate: Using children's childhood songs as literacy scaffolds in computer-based learning environments. Journal of Educational Computing Research 25(1), 17–34 (2001)
23. Hayes-Roth, B., Maldonado, H., Moraes, M.: Designing for diversity: Multi-cultural characters for a multi-cultural world. In: Proceedings of IMAGINA, pp. 207–225 (2002)
24. Moreno, K.N., Person, N.K., Adcock, A.B., Eck, R.N.V., Jackson, G.T., Marineau, J.C.: Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In: AAAI Symposium on Personalized Agents, Cape Cod, MA (2002)
25. Baylor, A.L., Kim, Y.: Simulating instructional roles through pedagogical agents. International Journal of Artificial Intelligence in Education 15(1), 95–115 (2005)
26. Cassell, J., Geraghty, K., Gonzalez, B., Borland, J.: Modeling culturally authentic style shifting with virtual peers. In: Proceedings of the 2009 International Conference on Multimodal Interfaces, pp. 135–142. ACM (2009)
27. Ogan, A., Finkelstein, S., Cassell, J.: Starting where the teachers are: Culturally-Sensitive Educational Technology Through a Teacher's Lens (in preparation)
28. Finkelstein, S., Scherer, S., Ogan, A., Morency, L.P., Cassell, J.: Investigating the influence of virtual peers as dialect models on students' prosodic inventory. In: Workshop on Child, Computer and Interaction (WOCCI 2012). ISCA, Oregon (2012)
29. Linn, M.C.: The knowledge integration perspective on learning and instruction. In: The Cambridge Handbook of the Learning Sciences, pp. 243–264 (2006)
30. McNeill, K.L., Krajcik, J.S.: Supporting Grade 5-8 Students in Constructing Explanations in Science: The Claim, Evidence, and Reasoning Framework for Talk and Writing. Pearson, Upper Saddle River (2011)
31. Chi, M.T.: Quantifying qualitative analyses of verbal data: A practical guide. The Journal of the Learning Sciences 6(3), 271–315 (1997)
32. Renn, J.: Acquiring style| The development of dialect shifting among African American children. Doctoral dissertation, The University of North Carolina at Chapel Hill (2010)
33. Schunk, D.H.: Peer models and children's behavioral change. Review of Educational Research 57(2), 149–174 (1987)