

# **Social Practice: Becoming Enculturated in Human-Computer Interaction**

Justine Cassell

Center for Technology & Social Behavior  
Northwestern University  
Frances Searle 2-148, 2240 Campus Drive. Evanston, 60208  
justine@northwestern.edu

**Abstract.** We present a new approach to the design, development and evaluation of embodied conversational agents (ECAs) that allows them to index identity through culturally and socially authentic verbal and non-verbal behaviors. This approach is illustrated with research we are carrying out with children who speak several dialects of American English, and the subsequent implementation and first evaluation of a virtual peer based on that research. Results suggest that issues of identity in ECAs are more complicated than previous approaches might suggest, and that ECAs themselves may play a role in understanding issues of identity and language use in ways that have promise for educational applications.

**Keywords:** Virtual Peers, Embodied Conversational Agents, Culture, Ethnicity.

The field of human-computer interaction (HCI) has its roots in Cognitive Science which, in turn, was born of Cognitive Psychology. For this reason, no doubt, the first several decades of research in HCI pursued a cognitive psychology-like “universalist” approach to the humans in the loop whereby all humans around the world were seen to be essentially identical in their behaviors – at least insofar as those behaviors might impact how they would interact with a computer – and all relevant behaviors were seen to be essentially universal. As an outcome of globalization – once systems designed in the US began to be shipped overseas, and once particular software interfaces were developed to allow communication among distributed work groups around the world – a “culturalist” approach began to take hold, whereby practitioners of HCI began to think about cultural differences. Today, still, when the topic of culture comes up in Human-Computer Interaction (HCI), the focus tends to be the culturalist one of differences between cultures, and how to characterize those differences. In this sense, the topic of culture in HCI falls under the rubric of user-modeling or personalization - how do we understand how people are, so that we can implement systems for different groups of people – in particular, people who are different from us. In this paper I am going to advocate a move away from a perspective that concentrates on differences and “how individuals are,” and towards a perspective where the focus is the production of difference in culturally-situated groups of people. This shift, derived from a reading of contemporary scholarship in Cultural Psychology, Anthropology, Sociolinguistics, and Education, is meant to push HCI researchers to reflect on

whether difference can be assumed to be *a priori* in any way, to be a trait of particular people, or rather whether difference is produced through social practices that take place in particular contexts. I will argue that it is the latter and, further, argue that HCI must therefore shift away from implementing different kinds of technologies for different kinds of people, and towards a paradigm where technologies are available to support the production of culture and identity.

The culturalist approach might be defined as assuming that “people from different cultures” are culturally different. In fact, “different cultures” really mostly means nation-states or countries, and so the assumption primarily concerns national character or national culture. In HCI, in part following the influential but problematic work by Hofstede [12] on globally-distributed teams, the topic of culture became a question of how to investigate the hidden differences between countries, and how to characterize them in ways that would allow us to understand the impact of those differences on the kinds of systems we built. As to the etiology of those differences, researchers from the culturalist tradition might explain that they arose out of the history of a given country, but would be hard-pressed to explain where the differences came from in individuals – there seemed to be an assumption that individuals were born different (or magically became different one day, when they hit puberty!), in a way that transcended the agency of those individuals. Today, outside of HCI, both the universalist and culturalist traditions have been increasingly supplanted by a socio-cultural approach to the study of culture [27], which is far less likely to accept the notion of “one underlying cultural logic or essence equally compelling to all members of the group” [13]. Instead, this contemporary scholarship sees people and culture as “mutually constituting” [24, 26]: (a) people are constructed through a history of engagement with particular cultural material (i.e., ideas, institutions, practices) - material that is associated with not just one, but a multiplicity of cultural worlds. And (b) culture itself is constructed through particular historical processes and maintained through the fact that those processes are reproduced everyday. That is, people re-create (and sometimes change) cultural worlds in the course of going about everyday business, and (c) the issue of who is defined as representative –or even a member – of a particular culture is never devoid of meaning. It often has to do with issues of power, and can erase majorities of citizens who for political, historical reasons are not seen as representative.

These postulates lead the sociocultural approach to problematize many of the concepts held dear by those who study cultural differences: problems with (i) the social context in which putatively differentiating behaviors are examined, (ii) intra-group variability whereby there may be as many differences within a particular country as across countries, (iii) the history and meaning of particular research participants’ experience with tasks that are used to identify their cultural behaviors, (iv) the effects introduced by the person asking the questions, (v) the history and connotation of the labels that are applied to particular behaviors (labels such as “trusting”, “independent”, “hierarchical”, “emotional”), and (vi) essentializing the characteristics of individual people, outside of their social interactions with others, on the basis of a group label, by assuming that people have some identifiable set of traits, characteristics, attributes or properties [20]. The solution proposed by sociocultural approaches is to see culture as something that is done rather than found, and to see culture as coming out of particular practices (people “live culturally” rather than “come from x

culture"). This perspective gives back agency to the populations we study, it makes identity more dynamic, and it allows us to understand the person in the person's cultural context.

One outcome of this perspective is an understanding that the concept of "difference" often hides the issue of who constitutes the norm. Thus, if we turn the lens on ourselves, we would have to admit that – while there are certainly regularities of behavior that characterize Americans – our knee-jerk conception of "American culture" may erase deep differences between ethnicities, social classes, gender, age, number of years since arrival in the US, and education-level (among others). We have to question whether the regularities we perceive are based on all Americans or one sub-group and, if one sub-group, who counts as the norm (Middle-class? Working class? African American? European American? Asian American? Hispanic? Arrived on the Mayflower? Arrived a year ago?). And I would argue that these issues arise for all countries/national cultures, whether or not they consider themselves relatively homogeneous or culturally-diverse.

Another outcome concerns the question of identity, and how identity plays out in communication and social interaction contexts. Culturalist perspectives often map particular ways of communicating onto particular cultures. "Japanese are hesitant in their speech", "Americans are confrontational", "poor people can't describe things adequately", "women's speech is emotional". In the sociocultural approach, on the contrary, identity constitutes the stuff out of which culture is made. In other words, identity can be seen as the demonstration, in a particular context, by a given person, of a set of behaviors and practices that index to others (both other members of the same group, and members of other groups) one's cultural community membership. Our communication with one another, then, always serves identity purposes – to establish our identity relative to one another. To give a personal example, while I rarely use the Brooklyn accent I was born with, I do occasionally find myself introducing Brooklyn features into my speech – when I am speaking to people with strong Brooklyn accents themselves, when I am trying to ally myself with non-academics or make a point about being a "down-home, normal" kind of person (as opposed to an ivory tower intellectual), or when I am ill or very tired. These examples demonstrate that I can deploy my accent (**not** consciously, for the most part, I might add) as ways of indexing my similarity to other people, or my adherence to certain cultural values, or my stance in a particular context (I'm tired, and not feeling like a grown-up professor).

One might argue that "Brooklynite" is hardly a cultural identity in the same way as American or Iraqi. But, I use my Brooklyn accent as an example because it illustrates a several essential points about indexing identity through communication. First, we are always participating in multiple cultural worlds [13]: ethnicity, gender, profession, country and region of origin, among many others. Our language may effect a kind of figure-ground process whereby particular cultural worlds are highlighted at a given moment. When looking at national cultural differences, the fact of our participation in multiple cultural worlds can sometimes become obscured, and yet I believe it is key to understanding how Culture (capital-C functions as well. My Brooklyn accent also demonstrates the important point that identities are not free of meaning in and of themselves. Difference usually entails power difference. Brooklyn is a low-status accent in most contexts (in fact, I speak the way I do now because of my parents'

desire to remove the Brooklyn from me; as they said, they wanted me to do well outside of Brooklyn!). This is true of accent, but it's also true of other aspects of speaking - deferring, hedges ("isn't it?"), and so forth. Ethnicity, social class, gender, region and country of origin are all aspects of cultural identity that are linked to ways of speaking that in turn are allied with differences in power and prestige.

If we belong to multiple cultural worlds, and are always signaling those identities through talk and behavior, then it means that the ways in which technologies look and talk also signal aspects of identity, whether their designers intend them or not. That is, users of technology constantly draw inferences about who the technology is, based on what it looks like, how it speaks, and how it acts, and those inferences lead to conclusions about the expertise of the technology, its intelligence, its value. Thus, for example, [22] found that GPS directions from a female voice are perceived as less accurate than those from a male voice.

If, in turn, the designers of technologies are trying to imbue them with cultural values they believe appropriate to the technology's users, then we must hope that the designers and users are on the same page about how cultural identity is produced. This problem is particularly keen for the technology known as an Embodied Conversational Agent or ECA [5], an interface to computing that has the form of a cartoon human on a screen, and engages in similarly human-like behaviors, with the goal of producing "natural interaction" between a computer and its user. The cartoon bodies force technology designers to make decisions about the visual traits the body displays (hair and eye color, skin tone, height, hair length, and so forth). And because the movements and speech of those bodies are driven off of underlying models, we are forced to make decisions about the nature of the ECA's behavior (the accent of speech, where gestures occur, the amount of mutual eye gaze, and so forth). The ECA was developed to highlight and take advantage of the essential role of the body in communication [5]; the embodied nature of the ECA means that its behaviors are interpreted within the context of our understanding of the role of the body in communication which, itself, is interpreted in the context of sociocultural understandings of interactions between people [4].



Fig. 1. Four Tutors [21]

In fact, in a spate of work that represents a laudable move away from the universalist position, where all ECAs were supposed to be "neutral representations" and nevertheless had the bodies and voices of white men and women, a number of researchers have recently addressed the issue of ethnicity in ECAs [19, 21]. However, the majority of these studies have modeled ethnic differences only through superficial surface traits such as skin color, hair style, clothing, etc. Additionally, the majority of these studies have been oriented towards the question of whether users

prefer to interact with a same or different ethnically-labeled agent, as opposed to the question of how they interpret the agents' identity, or how their own behavior changes when interacting with agents of putatively different cultural identities. Baylor [2], for example, implemented ECAs that varied in skin color and hairstyle, to represent Euro-American and African American agents. She found that African American

students reported preferring to interact with an African American agent; however, no actual evidence of increased learning outcomes was found. Similarly, Moreno and colleagues [21] implemented four ECAs to represent male and female African American and Euro-American math tutors. They found that whereas college students reported preferring to learn from a tutor of their own race and gender, their learning gains were greatest when the virtual tutor was a white male. Given the insights from sociocultural theory presented in the introduction, and the images from the study shown in Figure 1, this result is not so surprising. As can be seen, when gender and ethnicity were changed, a number of stereotypical features associated with gender and race were introduced. The professional white man and black woman contrast with a ditsy white woman and “homeboy” black man. These stereotypical racialized and gendered features may well have evoked unwanted inferences about intelligence and expertise that conflicted with the conception of the math tutor that the researchers intended. In one study that does apparently vary behavior, Maldonado and colleagues [19] implemented three agents that they describe as representing adolescent girls from Venezuela, Brazil and the United States in both look and action. No empirical evaluation was carried out with these agents, however, and no explanation is given of where the authors derived the traits and behaviors that they attribute to each ethnicity.

In contrast to the work cited above, which appears to treat racial and ethnic identity as a visually-apparent immutable division of peoples into groups, and where that division is not free from potentially noxious stereotypical attributions, in this paper we concern ourselves with identity and group membership as the demonstration in a particular context by a given person of a set of behaviors and practices that index to other members of the group, and to members of other groups, one’s cultural community membership. Ethnicity is one aspect of identity addressed here, but it is not the only one. In this more socioculturally-influenced approach to identity, physical appearance is not the only or even the most reliable index of ethnicity. We believe, on the contrary, that behavior may be preferable as an index of identity in ECAs, and that shallow visual stereotypes should be avoided unless the inferences about race and ethnicity evoked by those visual stereotypes is the subject of study. In addition, “identity” should be considered in the light of the different ways in which we present ourselves in various contexts.

Thus, we introduce the work below by reiterating that while language and nonverbal behavior are powerful cues to identity, most adults have mastery of more than one style. I don’t “speak Brooklyn” in front of a class, nor do I use the “professorial register” when hunting for bargains at a rummage sale. At every moment I am making choices about language use and behavior that index to my interlocutors my identity; my ability to seamlessly shift among several varieties is important for my professional and personal success. Some see the work of the classroom as instating the correct code in children – giving children the tools to succeed by teaching them the linguistic and behavioral codes that are associated with achievement in American schools. In practice, this approach often means that children are forbidden from using language varieties that are not those of white middle-class academic success. Others believe that educational practice must be explicitly linked to students’ everyday practices at home and amongst their peers [18]. And still others see the classroom as a place where hybridity and diversity (of language, but also of tools, roles, and social practices) should be nurtured [10]. The study reported here is a part of a larger research

program investigating the introduction of virtual peers (lifesize child-like ECAs) into classrooms and informal learning contexts as a way of thinking through the kinds of identity choices that children have and make in the classroom, during the development of an identity as student. In particular, in the current stage of this work, we are interested in the link between language difference, ethnic identity, and school literacy, and in the role of peers (as opposed to adults) in influencing children's choices about how (and who) to be in the classroom. The research reported here concentrates on the case of African American Vernacular English (AAVE). As background, in the US today, African Americans may speak a dialect of English known as AAVE, which has its own syntax, morphology and lexicon [9], and they also may use non-verbal communicative behaviors (such as eye gaze and gesture) in ways that differ from those used by Euro-Americans in similar situations [17]. Different dialects of AAVE exist in the United States, but many features remain constant across those different dialects. It is important to note that while some African Americans may use only AAVE forms in their speech, and use them in all contexts, others may use only some of the features of AAVE (the deletion of the copula, for example, so that they say "he running fast" rather than "he is running fast") and not use others. Still other speakers may employ some AAVE features in some social contexts, but their speech may be indistinguishable from GAE at other moments. And, of course, not only African Americans use AAVE features in their speech, as others may use AAVE to signal their identification with aspects of African American identity [8]. In addition, while the use of AAVE may be more prevalent in low-socioeconomic status communities, AAVE is certainly not restricted to members of that social class. Identity is thus constructed on the cultural level and on the micro-interpersonal level and the two are brought together in moment-by-moment behavior in conversation [16].

This link between language difference, ethnic identity, and academic achievement in African American children has been extensively studied, but the large body of research has resulted in little consensus. Some find a clear link only between teacher attitudes about language difference and academic achievement, such that teachers who believe that AAVE is a reflection of poor academic ability or intelligence are likely to find just what they expect [11]. Others find that instruction in the differences between General American English (GAE) and AAVE, and the contexts in which they are respectively welcomed leads to an improvement in literacy skills [28]. Still others argue that any effect of AAVE speaking on academic achievement is actually due to the fact that prevalence of AAVE features covaries with socioeconomic status, which in turn means that children may be more likely to attend schools with fewer resources, and of poor educational quality, as opposed to any causal effects of culture and dialect [14]. However research supporting the view that socioeconomic and dialect effects on learning can be separated also exists [7].

In the current research program we hope to add to the body of knowledge about the link between contextually-sensitive use of GAE and AAVE and academic achievement, but to abstract away from the instructional nature of current interventions where GAE forms are always taught by teachers in a classroom context, thereby adding a large measure of power to the equation [23]. After all, it is clear that social contextual factors, including the presence of code-switching peers, play a significant role in the development of code-switching skills, and the choices about language use that children make [6, 7]. To build a virtual peer capable of existing in a classroom where

both GAE and AAVE are used, then, we argue that the design process must be informed by careful research on actual communities of people in particular contexts, and must depend on a model that is derived from the verbal and non verbal behaviors that signal identity in those contexts. After each step of that research, the resulting virtual peer should be brought into classrooms so that we can assess the extent to which it is (a) accepted by children, (b) understood as indexing identity through language use, and (c) is capable of engaging in dialogues with children using one variety or other of language available to the children in that classroom.

To that end, we have been studying children in 7 schools and after-school programs in the Chicago area where AAVE is spoken. To date we have collected more than an hour of data from each of 44 third-grade children, ages 8 to 10 years, of whom 24 were African American and 20 were Euro American. In order to elicit different varieties of English, we introduced children to a number of different activities, with different conversational partners. The children were asked to describe pictures to AAVE and GAE-speaking unfamiliar adults. They were observed in interaction with a familiar GAE-speaking authority figure (the school principal), and they were observed in interaction in the hallways of their school with other children who spoke either AAVE or GAE. On the basis of this observation, we concluded that one of the African American children spoke GAE in all of the contexts that we observed, while one of the African American children code-switched from AAVE to GAE, depending on interlocutor. The other 22 children spoke AAVE in all the contexts that we observed. All 20 Euro American children spoke GAE throughout. The children were then divided into dyads by dialect use, and asked to imagine that people on one side of a raging river (represented by plastic figurines and a blue cloth) were trapped without any food or supplies (represented by weighted bags).



**Fig. 2.** Children building bridge

of children were asked to use a box of Legos to build a bridge that spanned the river and could support the family and the supplies so that the people could cross the river and bring back the food. They were told that the bridge must be long enough to span the river, strong enough to carry the weight, and wide enough to accommodate the figurines and the food. The researcher emphasized that they would have to test the structure with the weight, and then left the room. This activity was videotaped, and Figure 2 shows one of the camera views of one dyad. Once the children

were finished, the experimenter re-entered the room and explained that they would be allowed to bring their bridge into the main classroom and tell their teacher about how they had built the bridge. In order to prepare for this demonstration, the children were invited to take turns playing the roles of teacher and student.

The bridge-building exercise is designed to elicit peer-oriented language in an informal science task from the children, whereas the teacher – student exercise is designed to elicit use of formal conventions, science narrative, and code-switching from AAVE into GAE, if the children switch in formal school situations. Because the children are first engaged in the task with a peer and then, with that same peer, they

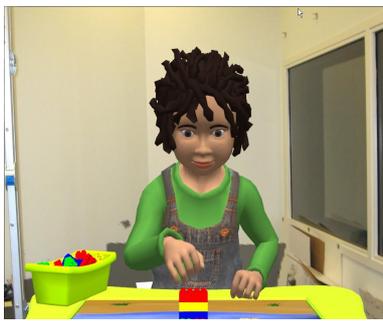
change to a classroom-oriented context, any changes in their behavior must be attributed to their sensitivity to the *context*, and not an outside influence that another person would introduce. And, because we recorded interactions of Euro American children as well, the task is standardized across both groups of speakers, allowing us to perform accurate comparisons.

Analyses of the children's language is still ongoing, but it is clear that both the GAE- and AAVE-speaking children made significant shifts in both their verbal and nonverbal behavior between the two tasks, in ways that index multiple community memberships. For example, while one AAVE-speaking child deleted the copula in his speech (a common feature of AAVE) in the peer task ("it a fit" instead of "it's a fit") he did not drop the copula when he was playing the role of teacher ("Why did you build your side of the bridge the way you built that?"). Likewise, while one dyad of GAE speakers used short commands with one another ("here – gimme that"), they switched to longer, less directive utterances when playing the role of teacher ("So, I heard that the first time it was strong and the second time it wasn't that strong, right?"). Table 1 shows the shift in the 22 AAVE features that constitute Craig and Washington's dialect density measure [7] between the bridge-building and teacher-student/classroom task. As should be clear, children reduced their use of AAVE substantially when speaking in the classroom context – both when playing the role of teacher and of student.

	<b>Bridge</b>	<b>Classroom</b>
Total AAVE Features	503	269
Total Words	5424	6297
Dialect Density	0.09	0.04

Nonverbal behavior was also different between the two tasks. One dyad of AAVE speakers completed the entire interaction without looking at one another – shifting their gaze between the Legos and the bridge. While playing the role of teacher and student, however, those same children looked at one another when asking or answering a question, and then shifted their gaze to the toys only when referring to them explicitly.

Child language scholars have long described register-shifting of this sort [1]. However, less studied is the link between register shifting and the acquisition of standard classroom talk. All but two of the AAVE-speaking children in our study spoke only AAVE in every context in which we were able to observe them. They were characterized to us by their



**Fig. 3.** Bridge-building virtual peer

teachers as monolingual AAVE and, further, as in need of remedial language teaching since they appeared incapable of learning GAE – the language their teachers require in class. Those same children, however, are clearly capable of using multiple kinds of language styles within a short period of time – including the use of GAE linguistic features – with a sophisticated notion of context as their guide. We must conclude, therefore, that their use of AAVE – and lack of use of GAE – in the classroom is just as much a choice and a way of indexing identity as is their employing the features of “teacher talk.”

The results described above will make up the behavior model that guides the actions of our virtual peer. We are currently analyzing the characteristics of science talk in the peer bridge-building and teacher tasks, so that the agent can also display a hybridity of classroom styles. The behavior model will be implemented in our virtual peer of ambiguous ethnicity (shown in Figure 3). In previous work [15] we have shown that ambiguous virtual peers may be implemented by using an iterative process of designing and showing to children until Euro-American and African American children are divided in opinion as to the ethnicity and gender of the agent. Maintaining an ethnicity-ambiguous *visual* appearance allows us to build an agent that can index identity fluidly through language use and nonverbal *behavior*, rather than being held to a rigid image of identity defined by the designer's choices of visual appearance.

The goal, as with our previous work on virtual peer support for proto-literacy behaviors [25], is to introduce the virtual peer into the classroom as a partner in explorations of the language choices that the child makes throughout the day. We do not misunderstand the nature of language use in the classroom – GAE is currently the language that the school will equate with high performance, and it currently seems valid to us to scaffold the child's development of that code – as long as it is seen as one choice among others – one aspect of the child's talk during a given day. Our prior work has demonstrated that, in fact, virtual peers are successful at influencing language choice – both in terms of scaffolding the use of decontextualized language [3] and in terms of scaffolding the use of GAE features by AAVE speakers [15].

In conclusion, note that a culturalist frame of reference would have obscured the ways in which identity and ethnicity are at play in the classroom, and the roles for technology therein. It is not the case that the African American children we observed speak and behave only in one particular culturally-specific way. Nor have these children acquiesced to adopting the linguistic trappings of achievement, by speaking GAE in the classroom even if they speak AAVE at home. Instead these putatively mono-dialectal AAVE-speaking children are able to use both AAVE and GAE features, as well as nonverbal behavior, to index the different identities at play during the school day — peers at work, and teachers and students in interaction. In order, then, to introduce to children a virtual peer capable of interacting around issues of code-switching, and to examine the different ways in which issues of language use might best be dealt with among peers in the classroom, I believe that we need to first build a virtual peer capable of indexing the multiple identities that the children recognize, and speaking the multiple dialects associated with them. As we have seen, language use and identity are not independent, and speaking in a particular way is a performance made up of choices taken from the arenas of grammar (including how one pronounces particular sounds, how one conjugates verbs, etc.), ways of speaking (including a sing-song voice, and the formality of one's language), and non-verbal behavior (including gestures, uses of the eyes and hands, shifts of the torso) – all ways of aligning with particular groups and particular social practices.

**Acknowledgments.** Thanks to Alberto Gonzalez, Kathleen Geraghty, Yolanda Rankin, and Francisco Iacobelli for their essential role in the work reported here, and to the other 2008-2009 fellows at the Center for Advanced Study in the Behavioral

Sciences for discussions and comments that vastly improved the manuscript. This work was made possible by grants from the National Science Foundation.

## References

1. Andersen, E.S.: *Speaking with Style: The Sociolinguistic Skills of Children*. Routledge/Kagan Paul, London (1990)
2. Baylor, A., Kim, Y.: The Role of Gender and Ethnicity in Pedagogical Agent Perception. In: *Proceedings of ELearn*, Phoenix, AZ (2003)
3. Cassell, J.: Towards a Model of Technology and Literacy Development: Story Listening Systems. *Journal of Applied Developmental Psychology* 25(1), 75–105 (2004)
4. Cassell, J.: Body Language: Lessons from the Near-Human. In: Riskin, J. (ed.) *Genesis Redux*, pp. 346–374. University of Chicago Press, Chicago (2007)
5. Cassell, J., et al.: *Embodied Conversational Agents*. MIT Press, Cambridge (2000)
6. Craig, H.K., et al.: Performance of elementary grade African American students on the Gray Oral Reading Tests. *Language, Speech, and Hearing Services in Schools* 35, 141–154 (2004)
7. Craig, H.K., Washington, J.A.: *Malik goes to school: Examining the language skills of African American students from preschool-5th grade*. Lawrence Erlbaum, New York (2005)
8. Cutler, C.: *Keepin' It Real: White Hip-Hoppers' Discourses of Language, Race, and Authenticity*. *Journal of Linguistic Anthropology* 13(2), 211–233 (2003)
9. Green, L.J.: *African American English: a linguistic introduction*. University of Cambridge Press, Cambridge (2002)
10. Gutiérrez, K.D., et al.: Rethinking Diversity: Hybridity and Hybrid Language Practices in the Third Space. *Mind, Culture, and Activity* 6(4), 286–303 (1999)
11. Hartwell, P.: Dialect Interference in Writing: A Critical View. *Research in the Teaching of English* 14(2), 101–118 (1980)
12. Hofstede, G.H.: *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival*. McGraw Hill, London (1991)
13. Holland, D.C.: *Identity and agency in cultural worlds*. Harvard University Press, Cambridge (1998)
14. Hunt, B.C.: Black Dialect and Third and Fourth Graders Performance on the Gray Oral Reading Test. *Reading Research Quarterly* 10(1), 103–123 (1974)
15. Iacobelli, F., Cassell, J.: Ethnic Identity and Engagement in Embodied Conversational Agents. In: *Proceedings of 7th International Conference on Intelligent Virtual Agents*, Paris, France, pp. 57–63 (2007)
16. Irvine, J.T.: Status and Style in Language. *Annual Review of Anthropology* 14, 577–581 (1985)
17. Johnson, K.R.: Black kinesics: Some non-verbal communication patterns in the Black culture. In: Samovar, L.A., Porter, R.E. (eds.) *Intercultural communication: A reader*, pp. 259–268. Wadsworth, Belmont (1976)
18. Lee, C.D., et al.: Every Shut Eye Ain't Sleep: Studying How People Live Culturally. *Educational Researcher* 32(5), 6–13 (2003)
19. Maldonado, H., Hayes-Roth, B.: Toward Cross-Cultural Believability in Character Design. In: Payr, S., Trappl, R. (eds.) *Agent Culture: Human-Agent Interaction in a Multicultural World*, pp. 143–176. Lawrence Erlbaum, New Jersey (2004)
20. Markus, H.R.: Pride, prejudice, and ambivalence: Toward a unified theory of race and ethnicity. *American Psychologist* 63(8), 651–670 (2008)

21. Moreno, K.N., et al.: Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In: Proceedings of 2002 AAAI Fall Symposium on Etiquette for Human Computer Work, Menlo Park, CA, pp. 77–80 (2002)
22. Nass, C.I., Brave, S.: *Wired for speech: how voice activates and advances the human-computer relationship*. MIT Press, Cambridge (2005)
23. Ogbu, J.U.: *Black American students in an affluent suburb: a study of academic disengagement*. L. Erlbaum, Mahwah (2003)
24. Rogoff, B.: *The Cultural Nature of Human Development*. Oxford University Press, New York (2003)
25. Ryokai, K., et al.: Virtual Peers as Partners in Storytelling and Literacy Learning. *Journal of Computer Assisted Learning* 19(2), 195–208 (2003)
26. Shweder, R.A.: Cultural Psychology: What Is It? In: Stigler, J.W., Shweder, R.A., Herdt, G. (eds.) *Cultural Psychology: The Chicago Symposia on Culture and Human Development*. Cambridge University, New York (1989)
27. Wertsch, J.V., et al.: *Sociocultural studies of mind*. Cambridge University Press, New York (1995)
28. Wheeler, R., Swords, R.: *Code-switching: teaching standard English in urban classrooms*. Urbana Ill: National Council of Teachers of English (2006)