UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Using Accent to Induce Cultural Frame-Switching

Permalink
https://escholarship.org/uc/item/90g319qn

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Dehghani, Morteza
Khooshabeh, Peter
Huang, Lixing
et al.

Publication Date
2012

Peer reviewed

Using Accent to Induce Cultural Frame-Switching

Morteza Dehghani, Peter Khooshabeh, Lixing Huang, Angela Nazarian, Jonathan Gratch

morteza,khooshabeh,lhuang,nazarian,gratch@ict.usc.edu
Institute for Creative Technologies
University of Southern California
12015 Waterfront Drive
Playa Vista, CA 90094

Abstract

Language plays a crucial role in the formation and categorization of one's ethnic identity. Recent work on linguistic accent emphasizes the role of accent in person perception and social categorization, demonstrating that accent serves as a meaningful ethnic category indicator. In this paper, we examine whether accent can be used to implement socio-cultural markers. We investigate whether the accent of a virtual character, as a marker for culture, can cause cultural frame-shifts in individuals. We report an experiment, performed among bicultural and monocultural individuals, in which we test the above hypothesis. Our results show that accent can have a socio-cultural effect on people's cognition.

Keywords: accent, culture, frame-switching, biculturalism

Introduction

There is a substantial amount of research focusing on the influence of language on human cognition. Language, at its core, is a cognitive tool that helps us construct shared meanings, categorize our knowledge, and thus piece together the various associative networks of collective experiences to form our perception of the world around us. According to the Whorfian hypothesis (Whorf, 1956), the grammar of a language and the manner in which the language is processed can provide a glance into the implicit organization of knowledge in that culture. This inherent influence is reflected in the belief that the grammatical structure of a language shapes our interactions due to the speakers inherent exposure to certain observations and mental representations, otherwise known as the principle of linguistic relativity (Chiu, Leung, & Kwan, 2007).

The results from previous studies with bilingual participants support the use of language as a means of activating associated cultural constructs. For instance, Chinese bilinguals randomly assigned to respond in Chinese reported more collective self-statements in open-ended self-descriptions, lower self-esteem, and more agreement with Chinese cultural views (Ross, Xun, & Wilson, 2002). These qualities are in line with traits commonly associated with Eastern collectivist cultures. The findings also suggest that cultural identities may be stored in separate knowledge structures and activated by the associated language. Marian and Neisser (2000) examined the retrieval of autobiographical memories in bicultural individuals and found that memories become more accessible when the language used at retrieval matched the language used at encoding. In other words, Russian-English bilinguals were able to remember more events from the Russian-speaking period of their lives when they responded in Russian and more experiences from the English-speaking period of their lives when they responded in English.

In their meta-analysis of cross-cultural studies of self-enhancement, Heine and Hamamura (2007) point out the limitations and confounds in existing research regarding whether language can independently prime a cultural mindframe. Lee, Oyserman, and Bond (2010) attended to these shortcomings in their study by exploring the effect of experimentally manipulated language use on value endorsements and self-descriptions. By unobtrusively manipulating language as a prime, they found that bilingual participants randomly assigned to use English reported significantly more self-enhancing social comparative judgements than those using Chinese. In addition, English users demonstrated more social distancing after under-performance and standardized their failure in the context of the behavior of their peers in order to protect their positive self-regard. This is in line with the expected behavior of someone from an individualist culture and touches on the work done by Tesser (2000) and Kwan et al. (2004) regarding the mechanisms employed to protect, augment, or maintain self-esteem. The above findings also support the claim that societies are socialized for both individual and collective mindsets, but responses are dependent on the cultural mindset salient at the moment of self-reflection (Lee et al., 2010).

This literature demonstrates the cognitive implications of experimentally manipulating language as a prime to prompt a shift in an individuals cultural mindset. The findings from these studies support the use of language as a marker for culture, however, in this paper we aim tease apart the specific aspects of language that act as cultural markers, e.g., by manipulating accents. Ethnolinguistic identity theory indicates that language plays a crucial role in the formation and categorization of one's ethnic identity (Giles, Bourhis, & Taylor, 1977). Rakič, Steffens, and Mummendey (2011a) compared the strength of vocal accents compared to facial visual appearance as indicators for ethnic categorization. They found that not only did accent serve as a meaningful ethnic category indicator, but participants also overlooked visual stimuli, such as facial features typically associated with a culture, in the presence of this more meaningful auditory information (Rakič et al., 2011a). In other words, sociopsychological processes were a more salient tool in determining memorability than objective stimulus features. In a later study, they examined the effect of accent workplace context. Job applicants with a standard German accent were compared to those with

nonstandard regional accents to test for biases influenced by the auditory information. Rakič, Steffens, and Mummendey (2011b) found that speakers with a standard German accent were perceived as more competent and more hirable than regional accent speakers. In addition, speakers with a Bavarian accent and the standard German accent were perceived as having higher socio-intellectual status than the remaining speakers. These results bolster the role of linguistic accent in person perception and social categorization.

In this paper, we examine whether accent can be used as a marker for culture by evoking cultural *frame-switching* (Hong, Morris, Chiu, & Benet-Martinez, 2000) in bicultural individuals. Cultural frame-switching refers to the idea that interpretive frames, in individuals who have internalized two cultural identities, can shift due to situational cues (Benet-Martinez, Leu, Lee, & Morris, 2002). For example, Hong et al. (Hong et al., 2000) prime Chinese-Americans with American, Chinese or neutral iconic images and demonstrate that participants in the Chinese prime condition interpreted the next task with more of a Chinese interpretive lens (focused more on external attributions) than those in the American or control conditions. On the other hand, participants in the American prime condition projected more American cultural values by focusing on individual attributions for the same task. A plausible explanation for frame-switching is that multi-cultural individuals have different norms and culturally specific systems of meaning (D'Andrade, 1984), which are activated depending on the context and the social environment. Hence, activation of different cultural interpretive frames can result in varied constraints on the individual's psychophysical behaviors.

The shift in the interpretive frames can be especially notable if there are significant differences in normative behavior between an individuals two cultures. There is growing evidence in the social sciences that culturally normative behaviors vary across cultures (see Henrich, Heine, & Norenzayan, 2010 for a review). In other words, behaviors that are considered typical in one culture may be considered abnormal in another. The variability of culturally normative behaviors and cognitions have been noted in different aspects of human behavior. Related to this work, researchers have shown cultural differences in morally motivated decision-making by identifying moral domains that are present (or salient) in some cultures but not in others (Shweder, Much, Mahapatra, & Park, 1997; Haidt, Koller, & Dias, 1993). Domains such as respect for authority and the saliency of the distinction between purity and impurity are some that have been identified in helping people characterize certain situations as morally tinged within one cultural group but not another. In this work, we carefully control for non-verbal behavior and manipulate only accent of spoken English, in order to experimentally evoke frame-switching and measure its effect on the participants' perception and cognition. Based on the previous literature, we predict individuals interacting with a counterpart having a culturally congruent accent will use moral frames specific to that

culture. Our method differs from studies that use language as a prime in the sense that in our experiment participants only listen and do not generate language explicitly.

Embodied Conversational Agents

Virtual agents, particularly when rendered as embodied conversational characters, are capable of providing a compelling multimedia platform that serves as an effective interface for research purposes, educational applications or entertainment. Embodied conversational agents (ECAs) make it possible to manipulate external features such as visual appearance, speech type, and contextual graphical environments. This ability makes ECAs a convenient platform to isolate unique cultural characteristics and realize them through simulation. Along with enhanced experimental control, ease of manipulations, consistency and controlled measurements (Loomis, Blascovich, & Beall, 1999), these features make ECAs useful and reliable tools for studying cultural cognitions. For example, we can objectively pinpoint certain social and behavioral characteristics that are relevant to specific cultures and implement them using virtual agent technology. There have been a small number of studies investigating how culturally congruent virtual agent characteristics can affect users' cognition. In an effort to examine the interaction between culture and ECA design in the domain of education, Rader, Echelbarger, and Cassell (2011) developed virtual peers that matched the dialect of children speaking African-American English and asked the children to complete a bridge building exercise. The children alternated playing the role of student and teacher as they explained the building process. Rader et al. found that students who tend to speak more dialected English did so less when they played the role of a teacher. This work suggests that the virtual peer and culturally congruent context, coupled with the role switch, influenced students to speak mainstream English, which is shown to be related to higher student achievement. In another line of work, Yin, Bickmore, and Cortes (2010) report that individuals who process information using peripheral cues perceived an agent tailored to their own culture as more persuasive and trustworthy.

Experiment

In this paper, we experimentally model frame-switching among bicultural individuals using the accent of an ECA and measure if their preference for certain moral acts are affected by this manipulation. Our hypothesis is that the accent of a virtual agent should affect people's perception of the culture of the agent. If this is true, then a virtual human that has an accent that is congruent with a participant's culture will elicit use of the congruent cultural frame. In order to test our hypothesis, we designed an experiment in which we control for non-verbal behavior of an ECA while manipulating only its accent. We recruited Iranian-American and American (US majoity culture) participants and had them read a story which included Iranian and US cultural values and customs. The participants were then asked to summarize the story and answer a few questions about the material.

Participants

Fifty-two Americans (mean age = 40) and Fourteen Iranian-Americans (Iranians living in Southern California for more than 5 years) (mean age = 34.61) participated in this study. The participants were recruited using craigslist.com and snowball sampling, which consisted of asking subjects to refer other subjects for the study. Each subject received $25 at the end of the experiment for participating. The participants were not aware that they were participating in a culture study. Each participant completed the task in individual experimental sessions.

Design

The study employed a between subject 2 X 2 full factorial design. The first factor is the culture of the participants (American or Iranian-American). The second factor is the agent's accent, which was either a standard American English or Iranian English accent, spoken by the virtual agent.

Stimuli

The participants were first asked to read a short story (Figure 2) about a student named Anthony who was asked to go to dinner at this classmate's (Shawn) house. After arriving in Shawn's house, Anthony comes to the conclusion that Shawn's parents were not expecting a guest for dinner. Anthony complements Shawn's dad about an art piece in the house, and Shawn's dad insists that Anthony should take the picture. Finally, Anthony's friend picks him up from Shawn's house before dinner was served. The story included a balanced number of American and Iranian cultural products (e.g. proverbs), values, and events that could be tied to the celebration of Iranian New Year and to Saint Patrick's day[1]. None of these idea units were explicitly labeled with their cultural referent (there was no explicit reference to St. Patrick's day as such) and the idea units were interleaved so as to minimize memory distortions due to recency or primacy. After reading the story, participants interacted with an ECA.

Rapport Agent

The agent used in this experiment, Utah (Hartholt, Gratch, Weiss, & Team, 2009) (Figure 1), is designed to establish rapport with human participants by providing contingent feedback while a user is speaking. To produce feedback, the agent first detects and analyzes in real-time the human speakers' audiovisual features, which are silence, head nod, eye-gaze (looking at the agent or not) and smile. The audio feature detector extracts intensity from the raw signal every 100ms using Praat[2]. With the intensity information, it outputs a binary feature, speech or silence, every 100ms. The visual feature detector[3] tracks the position of the face, the facial feature points, the direction of eye gaze and the smile level. With this information, it outputs visual features indicating whether the
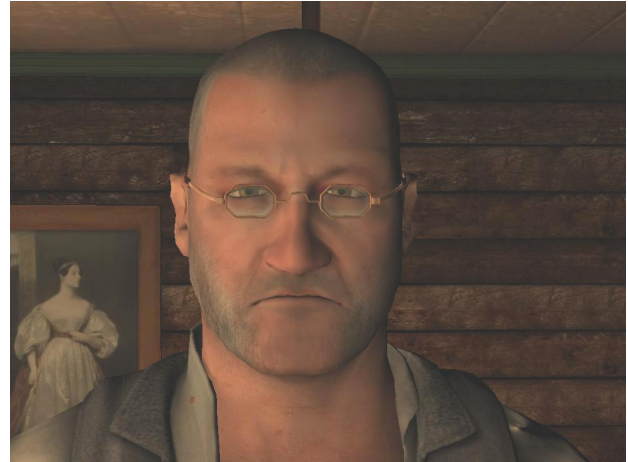

Figure 1: The ECA used in our Experiments

human is nodding or not, looking away or not, and smiling or not. Based on the perceived audiovisual features, the response model (Huang, Morency, & Gratch, 2011) decides, in real-time, the most appropriate responses, such as head nod and smile. These different styles of animations are first converted into Behavior Markup Language (BML) (Kopp et al., 2006) and then sent to an action scheduler, which keeps track of the duration of each animation. If the current animation has not been completed, the new animation will be ignored. The BMLs are passed to Smartbody (Thiebaux & Marsella, 2007), a virtual human animation system designed to seamlessly blend animations and procedural behaviors. Finally, the byproducts of Smartbody are rendered by a commercial game engine, Gamebryo[4], and displayed to users. For the experiment, the voice of the ECA was prerecorded using the voice of the second author, whom is familiar with both Iranian and American cultures. The second author did not participate in recruiting nor in running the experiments.

Procedure

After participants finished reading the story, the virtual agent greeted them, explained an overview of the research center and asked them to verbally summarize the story they had just read. Next, they filled out a questionnaire about the appropriateness of certain actions and intentions of the characters within the story. They were specifically asked the following two questions: 1. Was it appropriate for Anthony to leave before dinner? 2. Do you think Shawn's parents really wanted to give the picture to him? Our hypothesis predicts that participants should use culturally congruent frames to interpret and answer these moral questions. The Iranian cultural frame suggests that it is not appropriate to refuse someone's generosity and hospitality. If cultural frame-shifting does indeed take place for Iranians-Americans when interacting with the culturally congruent agent, then they should say it is inappropriate for Anthony to leave dinner early. For the second

---

[1]Saint Patrick's day is widely celebrated in North America

[2]Praat, http://www.fon.hum.uva.nl/praat/

[3]OKAO Technology, http://www.omron.com/r_d/coretech/vision/okao.html

[4]http://www.gamebryo.com

Figure 2: Excerpt from the story read by participants

question, the Iranian frame could interpret the event as an instance of Iranian hospitality, especially when it comes to sharing their cultural artifacts (in this case Persian miniature). Next, participants were asked several questions about different emotions of the characters in the story. Lastly, to check the effectiveness of our manipulation, participants answered the following two questions in a random order: 3. Did the agent have more of an American accent or Middle-Eastern accent? 4. Did the agent appear more Western or more Middle-Eastern? Each question was answered on a 6-point scale (1 = No he did not, 6 = Yes he did; 1 = Not at all appropriate, 6 = Completely appropriate; 1 = Very much American, 6 = Very much Middle-Eastern; 1 = Very much Western, 6 = Very much Middle-Eastern).

Results

For both manipulation check questions, we used the responses to questions 3 and 4 as dependent variables in a 2 X 2 ANOVA, where the first factor was the culture of the participants (American or Iranian-American) and the second factor was the accent of the agent (American or Iranian). There was a main effect of agents' accent for both questions (appearance: $F(1, 62) = 11.038$, $p = 0.0015$; accent: $F(1, 62) = 68.1434$, $p < 0.001$). The agent with an Iranian accent was viewed to not only have a more Middle-Eastern accent but also appeared more Middle-Eastern. Also, there was a main effect of culture for the appearance question ($F(1, 62) = 4.276$, $p = 0.0428$) where Americans ranked the agent as more Middle-Eastern looking than did the Iranian-Americans. The responses to the first question were used as the dependent variable in a 2 X 2 ANOVA, with similar factors as above. There was a significant interaction between the two factors ($F(1,62) = 4.3649$, $p = 0.0408$). A planned comparison revealed that Iranian-American participants who interacted with the agent with an American accent viewed Anthony leaving before dinner as more appropriate than the Iranian-Americans who interacted with the agent with an Ira-

nian accent (Welch Two Sample one-tailed t-test: $t(10.596) = 10.596$, $p = 0.0512$)[5] (Figure 3). However, this difference did not reach significance for Americans ($t(49.48) = 1.4179$, $p = 0.1625$).

Next, using the responses to the second question as the dependent variable, we ran the same 2 X 2 ANOVA. There was a main effect of culture ($F(1, 62) = 9.5759$, $p = 0.0029$), where Iranian-Americans ranked this question lower than Americans. There was an interaction between culture and accent of the agent ($F(1, 62) = 9.5759$, $p = 0.0029$), where Iranians who interacted with the American accent agent indicated that Shawn's parents didn't want to give the picture to Anthony compared to those in the Iranian accent condition (Welch Two Sample t-test: $t(9.008) = 3.1186$, $p = 0.0123$). However this difference did not reach significance for Americans ($t(49.983) = -1.2908$, $p = 0.2027$).

There were also interactions between culture and agent's accent for Anthony's feeling of happiness ($F(1, 62) = 5.0474$, $p = 0.0282$) and satisfaction ($F(1, 62) = 12.7468$, $p < 0.001$). For happiness, Iranian-Americans interacting with the agent with an Iranian accent rated Anthony's happiness higher than those Iranian-Americans who interacted with an American accented agent ($t(11.725) = 2.4057$, $p = 0.0336$), and vice versa for American participants ($t(49.408) = -2.2325$, $p = 0.0301$). The same significant trend held for Anthony's satisfaction (Iranian-Americans: $t(11.855) = 3.2666$, $p = 0.0068$); Americans: $t(48.815) = -3.543$, $p < 0.001$). Similarly, there was a two-way interaction between culture and accent for Shawn's parents' satisfaction ($F(1, 62) = 7.0429$, $p = 0.0101$) and the effect also approached significance for Shawn's satisfaction ($F(1, 62) = 3.0044$, $p = 0.0880$). Similar to the emotions reported above, participants ranked the satisfaction of Shawn's parents and Shawn higher when the accent

---

[5]A power test revealed that if we had the same number of Iranian-American participants as American participants, with probability of 99.15% a two-tailed test with the means of the above sample would have reached significance.
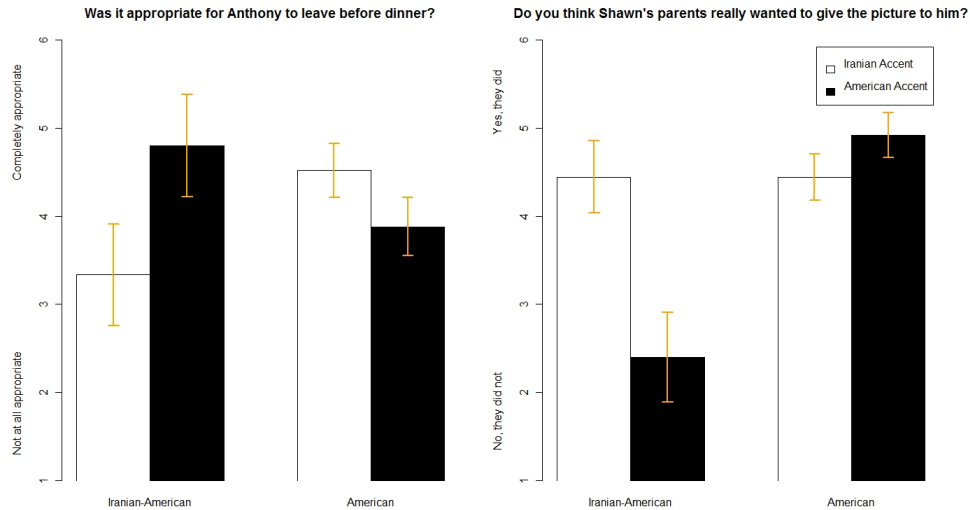
Was it appropriate for Anthony to leave before dinner?

Do you think Shawn's parents really wanted to give the picture to him?

Figure 3: Interaction between culture and accent of the agent

of the agent matched their own accent (Shawn's parents': Iranian-Americans: t(10.121) = 3.6028, p = 0.0047; Americans: t(49.362) = -1.4053, p = 0.1662, Shawn's: Iranian-Americans: t(5.953) = 0.7476, p = n.s.; Americans: t(49.054) = -2.2229, p = 0.0309).

Discussion

This study provides evidence of how spoken accent can make a socio-cultural effect on people's cognition. In a fully factorial design, Iranian-Americans who interacted with a virtual agent that spoke Middle Eastern accented English were more likely to use a congruent cultural frame to interpret a morally charged scenario. The compelling aspect of this effect is that the accented virtual agent's visual appearance was identical across experimental conditions and the only manipulation was the agent's accent. Although simple effects were not significant for the Americans, the trends were in the correct direction. The trends could be due to the fact that most of the American participants recruited in this study have a multicultural background (33% African-American, 14% Latino).

To our surprise, our manipulation also affected people's evaluations of the emotions of the characters in the story. We speculate that this effect might be due to the fact that participants who interacted with an agent whose accent matched their own attributed intentions and goals that are congruent to their own cultural mindset to the characters and, as a result, appraised the situation for the characters more positively. Therefore, when Iranian-Americans interacted with the ECA that had an Iranian accent, they appraised the intentions and goals of the characters to be more in line with the Iranian culture, hence they evaluated the situation more positively, compared to Iranian-Americans who interacted with the American accented ECA.

We acknowledge the low number of Iranian-Americans

participants in our study. However, we would like to note that the probability of replication of a result is dependent on p-levels but not affected by sample size (e.g. Killeen, 2005).

In summary, contributions of this work are two-fold. First, the study adds to the literature in cross-cultural psychology by showing that just spoken language accent can induce cultural frame shifts. Second, this work makes a methodological contribution to the field of human-computer interaction and experimental psychology. Our results also have implications for teaching cross-cultural fluency and competency.

References

Benet-Martinez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating biculturalism: Cultural frame switching in biculturals with oppositional versus compatible cultural identities. *Journal of Cross-Cultural Psychology*, *33 (5)*, 492-516.

Chiu, C., Leung, A. K., & Kwan, L. (2007). The handbook of cultural psychology. In S. Kitayama & D. Cohen (Eds.), (p. 259-276). Guilford Press.

D'Andrade, R. (1984). Cultural meaning systems. In

R. Schweder & R. Levine (Eds.), *Culture theory. essays on mind, self and emotion* (p. 88-119). Cambridge: Cambridge University Press.

Giles, H., Bourhis, R., & Taylor, D. M. (1977). Language, ethnicity and intergroup relations. In H. Giles (Ed.), (p. 307-348). Academic Press.

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*, 613-628.

Hartholt, A., Gratch, J., Weiss, L., & Team, T. G. (2009). At the virtual frontier: Introducing gunslinger, a multi-character, mixed-reality, story-driven experience. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. Vilhjálmsson (Eds.), *Intelligent virtual agents, lecture notes in computer science* (p. 500-501). Springer.

Heine, S. J., & Hamamura, T. (2007). In search of east asian self-enhancement. *Personality and Social Psychology Review*, *11 (1)*, 4-27.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61-83.

Hong, Y., Morris, M. W., Chiu, C., & Benet-Martinez, V. (2000). Multicultural minds. a dynamic constructivist approach to culture and cognition. *American Psychologist*, *55 (7)*, 709-20.

Huang, L., Morency, L. P., & Gratch, J. (2011). Virtual rapport 2.0. In *International conference on intelligent virtual agents.*

Killeen, P. R. (2005). An alternative to null-hypothesis significance testing. *Psychological Science*, *16*, 345-353.

Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., . . . Vilhjlmsson, H. (2006). Towards a common framework for multimodal generation: The behavior markup language. In *International conference on intelligent virtual agents* (pp. 21–23).

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review*, *111 (1)*, 94-110.

Lee, S., Oyserman, D., & Bond, M. H. (2010). Am i doing better than you? that depends on whether you ask me in english or chinese: Self-enhancement effects of language as a cultural mindset prime. *Journal of Experimental Social Psychology*, *46*, 785-791.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior research methods instruments computers a journal of the Psychonomic Society Inc*, *31*(4), 557–64.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, *129 (3)*, 361-368.

Rader, E., Echelbarger, M., & Cassell, J. (2011). Brick by brick. In *Proceedings of the 2011 annual conference on human factors in computing systems - CHI '11* (p. 2971).

Vancouver, BC, Canada. doi: 10.1145/1978942.1979382

Rakič, T., Steffens, M. C., & Mummendey, A. (2011a). Blinded by the accent! the minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100 (1)*, 16-29.

Rakič, T., Steffens, M. C., & Mummendey, A. (2011b). When it matters how you pronounce it: The influence of regional accents on job interview outcome. *British Journal of Psychology*, *102 (4)*, 868-883.

Ross, M., Xun, W. Q. E., & Wilson, A. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin*, *28 (8)*, 1040-1050.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). Morality and health. In A. Brandt & P. Rozin (Eds.), (p. 119-169). New York: Routledge.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4 (4)*, 290-299.

Thiebaux, M., & Marsella, S. (2007). Smartbody: Behavior realization for embodied conversational agents. In *In 7th international conference on intelligent virtual agents (iva).*

Whorf, B. (1956). *Language, thought, and reality*. Cambridge, MA: Technology Press of MIT.

Yin, L., Bickmore, T. W., & Cortes, D. E. (2010). The impact of linguistic and cultural congruity on persuasion by conversational agents. In *Proceedings of IVA, lecture notes in computer science* (pp. 343–349). Philadelphia, PA: Springer.