A Verbal and Gestural Corpus of Story Retellings
to an Expressive Embodied Virtual Character

Jackson Tolins[1], Kris Liu[1], Michael Neff[3], Marilyn Walker[2], Jean E. Fox Tree[1]

[1]University of California, Santa Cruz, Department of Psychology
[2]University of California, Santa Cruz, Department of Computer Science
[3]University of California, Davis, Department of Computer Science
E-mail: jtolins@ucsc.edu, kyliu@ucsc.edu, foxtree@ucsc.edu, ma, M@ucsc.edu, mpneff@ucdavis.edu

Abstract

We present a corpus of 44 human-agent verbal and gestural story retellings designed to explore whether humans would gesturally entrain to an embodied intelligent virtual agent. We used a novel data collection method where an agent presented story components in installments, which the human would then retell to the agent. At the end of the installments, the human would then retell the embodied animated agent the story as a whole. This method was designed to allow us to observe whether changes in the agent's gestural behavior would result in human gestural changes. The agent modified its gestures over the course of the story, by starting out the first installment with gestural behaviors designed to manifest extraversion, and slowly modifying gestures to express introversion over time, or the reverse. The corpus contains the verbal and gestural transcripts of the human story retellings. The gestures were coded for type, handedness, temporal structure, spatial extent, and the degree to which the participants' gestures match those produced by the agent. The corpus illustrates the variation in expressive behaviors produced by users interacting with embodied virtual characters, and the degree to which their gestures were influenced by the agent's dynamic changes in personality-based expressive style.

Keywords: spontaneous communication, gesture production, virtual agents, personality, stylistic adaptation, gestural and linguistic paraphrase

1. Background and Motivation

We present a corpus of 44 human-virtual agent story retellings that were designed to collect both verbal and nonverbal aspects of human-agent interaction in order to investigate whether people would gesturally entrain to a virtual agent. We used a novel method involving retelling a story in installments in order to investigate the effect on human gestural and verbal behavior of a protocol in which the intelligent embodied agent changed its gestural expressive style over time, either starting as introverted and becoming extraverted, or vice versa.

Previous research suggests that intelligent virtual agents will be viewed more positively if they are both capable of nonverbal expressive behavior, and capable of adapting their behavior to their conversational partner. This adaptation may be based on personality, specific behaviors of the partner, emotional expressiveness or matching, or cultural norms (Andre et al 2014, Endrass et al, 2010; Kopp et al 2006, Mairesse & Walker 2010, 2011; Salem et al, 2012, Hartmann et al 2005). Some studies have suggested that agents that can behaviorally coordinate will be viewed more positively, are more believable, and are more persuasive (Bailenson, et al., 2008, Tapus & Mataric, 2008, Andre et al 2000). People respond to coordinating computers in ways that are similar to how they respond to coordinating people. If the computer mimics their own verbal personality style, people like the computer more (Nass & Moon, 2000). People also tend to have a preference for interacting with agents whose nonverbal expressive behaviors present a personality that matches their own (Bailenson & Yee, 2005; Liu et al., 2013, in press). However, people do seem to be forgiving of machines in ways that some

researchers have found they are not of each other (Aronson & Linder, 1965). Computers' changes to be unlike humans did not result in people disliking the computer (Nass & Moon, 2000). Investigation in human-agent interaction has explored both how humans adapt to agents (Branigan, et al., 2003; Heyselaar et al., 2014), and how agents can be designed to be perceptive towards users' communicative behaviors and adaptive in response (Buschmeier et al., 2009; De Jong, et al., 2008; Walker, et al., 2007, Mairesse & Walker 2010, 2011).

While studies have considered how humans adapt to computers linguistically, through the repetition of syntactic constructions or register, few studies have considered the adaptation of nonverbal expressive behavior in relation to embodied conversational agents (for an exception see Kramer et al., 2007). Similarly, while researchers have attempted to implement systems that imbue agents with automatic alignment between the perception and production of gestures, less work has been done on user gestures directed towards agents in the context of interaction. Currently, no studies have investigated the degree to which a person's gesturing may be influenced by the presence and expressive style of an agent's gestures. People interacting with a hand-held virtual agent do produce gesture and other nonverbal displays (Bickmore, 2002), although they produce fewer gestures, gaze patterns away, and head nods in comparison to interactions with a human interlocutor. Given the similar reduction in gesturing in human-human conversations without visual co-presence, for example when speaking on the telephone (Bavelas, et al., 2008), it is unclear if the gestures produced by the participants interacting with the handheld virtual agent were indeed communicative, or were rather gestures produced to facilitate speech production (Rauscher, et

al., 1996). That some participants in this study gestured with both hands, including the one holding the virtual agent device, and that many produced gestures that would be out of the visual field of the agent, suggests that it is not necessarily a foregone conclusion that people will gesture communicatively towards agents. Importantly, for the development of agents capable of perceiving and adapting to the nonverbal expressive behaviors of users, these users should gesture communicatively towards the agents. Similarly, parallel to the review presented above, it is likely that in a communicative interaction, users may adjust their gestures to match those of the agents.

By collecting participants' gestures as they interact with a virtual agent, we are able to capture both direct repetitions of single gestures as well as general stylistic adaptation in expressive behavior. In interactive dialogue, speakers repeat each other's gestures across conversational turns (Holler & Wilkin, 2011; Mol, Krahmer, Maes, & Swerts, 2012). Beyond one-to-one repetition of single gestures, speakers also adapt to the stylistic expressive behavior of their conversational partner, both in speech (Giles, Coupland, & Coupland, 1991) and in gesture (Bergmann & Kopp, 2012; Tolins, Liu, Wang, Fox Tree, Neff, & Walker, 2013, 2016).

Stylistic expressive behavior may be particularly useful as a cue to adaption in that it can be taken to reflect the personality of the speaker. A number of studies have explored nonverbal expressive correlates of personality, both in humans and in conversational agents. The personality dimension of extraversion in particular is highly visible, allowing for accurate perception and judgment (Funder & Dobroth, 1987; John & Robins, 1993). Speakers who score high on scales of extraversion tend to produce broader gestures that are further from the body (Lippa, 1998; Riggio & Friedman, 1986). In addition, extraverts gesture more frequently and with more rapid movements than introverts (LaFrance, Heisel, & Beatty, 2004). Manipulations of these particular dimensions of gesture have been incorporated into virtual agents (e.g. Hu, Walker, Neff, & Fox Tree, 2015; Neff, Wang, Abbott, & Walker, 2010). People observing these agents accurately interpret the expressive style visible in the gestures as indications of the agents' personality (Neff et al., 2010; Liu, Tolins, Fox Tree, Neff, & Walker, 2013, 2016).

We present a corpus of human-agent verbal and gestural story retellings. We used a novel data collection method where an agent presented story components in parts for a person to retell, concluding with the person retelling the story as a whole. This method allows observation of human behavioral changes while an agent changes behavior over time. As adaptation towards a conversational partner has been correlated with both affiliation (Chartrand & Bargh, 1999) as well as conversational success (Louwerse, Dale, Bard, & Jeuniaux, 2012), willingness to adapt towards a conversational virtual agent may be a critical part of usability, both in terms of the user adapting towards the agent and the agent changing its expressive behavior to adapt towards the user. In our study, we manipulated whether the agent moved from nonverbal expressive correlates of extraversion to introversion, or from introversion to extraversion. The corpus is available at nlds.soe.ucsc.edu/corpora.

2. Corpus Description

We collected a corpus of verbal and nonverbal communication between participants and an agent whose expressive behavior changed over time, from extraversion to introversion, or from introversion to extraversion. The corpus provides speech transcription as well as information about the frequency and variability in nonverbal communication towards agents. This allows exploration of communicative behaviors directed towards agents and the degree that people adapt their gestures to those of an agent. The corpus also includes information about the retellers' personalities.

2.1 Participants

Forty-four participants were drawn from the University of California, Santa Cruz research participant pool. Participants were compensated with partial course credit.

2.2 Agent and Agent's Expressive Behaviors

Two interactive story scripts were created. The narratives were drawn from a corpus of spontaneously produced near-miss stories told by University of California, Santa Cruz undergraduates in 2014. Near-miss stories demonstrate typical narrative arcs, including a set-up, climax, and resolution and include topic matter relatable to by the participant population. The two selected stories were transcribed and the transcriptions synthesized using AT&T text-to-speech, using the American English voice Crystal (AT&T Natural Voices SDK) with minor adjustments and changes towards understandability. Each story transcription was separated into five separate story installments. A transcript of one of the snippets follows:

> So I spun off and it was like, it was like really smoky, so I thought that the car was going to blow up, so I don't know, I guess like, adrenaline had like kicked in or whatever and I couldn't get out of my door 'cause it was all bent up, but I crawled out of the passenger seat.
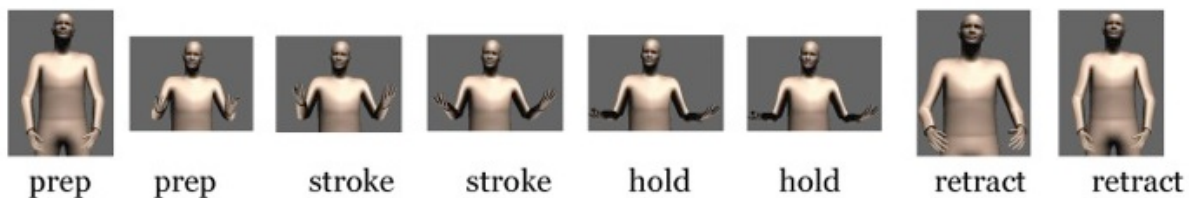
Figure 1: Phases of the gestures produced by the agent, synchronized to story events.

Agent gestures were scripted to co-occur with the synthesized speech. Noun phrases and verb phrases from the transcript were selected for co-produced gestures. The agent's gestures were drawn from a library of previously collected motion capture data. Gestures were selected such that they represented meaningful aspects of the utterance (McNeill, 2005). As such, the gestures were either iconic or metaphoric (McNeill, 1992; 2005). Iconic gestures represent physical, or concrete, aspects of the content of a message, whereas metaphoric gestures are used to represent abstract content. Previous research has suggested that speakers will repeat content-rich gestures (Holler & Wilkin, 2011; Kimbara, 2008). By restricting the agent's gestures to those that may be interpreted as a meaningful part of the agent's message, we hoped to be able to quantify the degree to which participants displayed adaption through gesture repetition.

Gestures were scripted to include *preparation* phases, in which the agent's hands were drawn up into the gesture space from rest, *strokes*, the meaningful part of the gesture, and *retractions*, in which the gestures were brought out of the gesture space. See Fig. 1. A number of gestures included a *post-stroke hold*, in which the hands were held in place. This was particularly true for instances in which two gestures occurred close together: Rather than retracting the hands after each gesture, the agent would hold between strokes, mimicking a larger gestural utterance as would be produced in human communication (Kipp, Neff, Kipp, & Albrecht, 2007).

The expressivity of the gestures was manipulated so that the agent expressed varying degrees of extraversion/introversion across story parts. See Fig. 2.
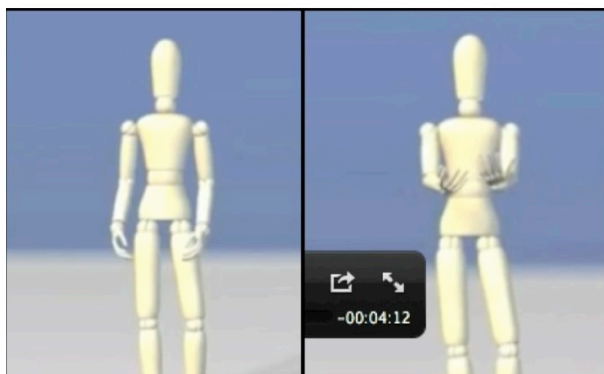
The expressive profile was manipulated such that each story part was told with a distinct level of gestural extraversion using a sliding scale representing five levels of extraversion/introversion as expressed through the nonverbal expressive behavior. Two versions of each story were created. In one the agent moved from extraverted to introverted gesturing behavior. In the other, the agent moved from introverted to extraverted. A number of coordinated dimensions were manipulated, including spatial extent, the distance of the agent's hands from the center of its body, and gesture rate, with gestures removed from the original script to correspond to positions along the sliding scale. Both expansiveness of gestures and gesture rate have been positively correlated with high levels of extraversion in humans (LaFrance, Heisel, & Beatty, 2004) and agents (Liu et al., 2016). The most extraverted agent thus had the largest, fastest gestures, with a high number of different gestures throughout the story. The most introverted agent gestured much less frequently during the same stretches of talk, and the gestures they did produce were smaller and slower. Introverted agents also displayed less body movement. For details, see Neff et al. (2010) and Hu et al (2015, 2016). These different gesture profiles were produced in tandem with the same script, allowing for an exploration of the role of nonverbal behaviors alone.
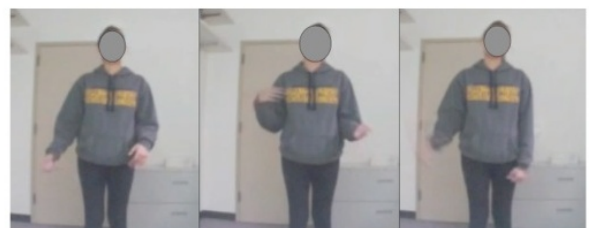

Figure 3: Human Retelling of a Story installment back to the agent.

The agent was rendered using a 3D wooden manikin animation with five-fingered hands. Manufactured gestures were combined into a single scripted file and rendered into a movie using the animation software Maya (Autodesk). The rendered videos presented the entire model of the agent facing directly towards the participant (Figure 2). The agent was projected using a Ricoh Ultra Short Throw Projector (www.ricoh-usa.com), which displayed the agent at roughly eye level and with a height of four feet and a maximum arm span of three feet. This setup was designed to make the human subject feel as though they were talking to another life-sized person. The humans

Figure 2: Agent varying gesture performance to manifest personality on the extraversion scale .

stood up in a small enclosed space while talking to the agent during the story telling and retelling. This elicited realistic gestural performances from the humans. See Fig. 3.

2.3 Procedure

Each participant engaged in two different story retelling interactions with the virtual agent. For one story, the first installment of the agent's nonverbal expressive behavior was highly extraverted and then story installments became introverted over the course of the story, and for the other story the agent's initial story installment was designed to be highly introverted, and then the protocol was for the agent to become more and more extraverted over each story installment. The order of the two stories and the direction of the agent's adaptation were counter-balanced across participants.

Participants were informed that the purpose of the study was to interact with a virtual agent, the Story-Tron 9000, designed to learn and tell stories. The interactivity of the agent was emphasized by pointing at the cameras and microphones through which the participants were told the agent saw and heard the participant, and by making the participant wave and say hello to activate the agent at the beginning of the story. In fact, no participant behavior was necessary for the agent to respond; a research assistant went into another room and pressed *play* on the video. Each story was told in five parts of equivalent length, with the agent prompting the participant to repeat back each part. After the repetition of the fifth story part, the agent prompted the participant to retell the story in its entirety. After each story, participants completed a short survey on how they perceived the agent and the interaction. Following the second of these post-interaction surveys, the participants also completed a general questionnaire, as well as the TIPI personality survey (Gosling et al., 2003).

2.4 Transcription and Annotation

Participants were recorded through three small cameras aligned at 0, 45, and 90 degrees relative to the participants' orientation towards the virtual agent. The three feeds of the participant, along with the projection of the virtual agent and the speech recording, were aligned through CaptureSync (http://www.bensoftware .com/capturesync/;see Figure 2). These composite recordings were used for the transcription of talk and gestures by trained coders using the transcription software Elan (Brugman & Russel, 2004; http://tla.mpi.nl/tools/tla-tools/elan/). Coders were trained on pilot data and then individually transcribed the talk. A transcript of one of the agent snippets (repeated here for comparison) and three human retellings follows:

Agent: So I spun off and it was like, it was like really smoky, so I thought that the car was going to blow up, so I don't know, I guess like, adrenaline had like kicked in or whatever and I couldn't get out of my door 'cause it was all bent up, but I crawled out of the passenger seat.

Retelling 1: And I spun out and it was really smoky and I thought my car was going to blow up and like adrenaline hit again and I tried to get out of my car but I couldn't because it was so bent up that I crawled out of the passenger seat.

Retelling 2: The car hit me and I was spinning out of control and uh the car started smoking and I didn't know what to do I couldn't go out of my door so I went out the passenger's seat.

Retelling 3: Yeah so you um you were um your car was all bent up and or you were spinning and you were like freaked out and you wanted to get out of your car and adrenaline kicked in but you couldn't get out of your car because your car door was open so you had to crawl out the passenger's window.

These examples show similarity in some retelling choices, such as mentioning the concept of spinning, and differences in other retelling choices, such as retelling in first person or second person.

Coders then coded the gestures using a tiered structure. These transcriptions were then double-checked by a second coder and modified as needed. Gesture phase and phrases were coded (Kita et al., 1998), capturing the temporal structure of the gestures. Each gesture consisted of at least a stroke phase, and may or may not have also contained a preparatory phase, a hold phase, and a retraction phase. These phases were combined into larger gesture units called phrases, with each phrase coded for type as well as the lexical items with which it co-occurred. Gestures were categorized as iconic, metaphoric, deictic, and beat. The spatial extent of the gestures was coded using a previously developed scheme that captures height, distance from the body, and radial orientation (Kipp et al., 2006; Kipp, Neff, & Albrecht, 2007). Based on the words produced with gestures, participants' gestures were matched to those produced by the agent in regards to three features: *a*. location – whether the gesture co-occurred with the same words that the agent produced gesture with, *b*. form – whether the participants' gesture matched the form of the agent's gesture at this same location, and *c*. handedness – whether the participant expressed the gesture using the same handedness as originally produced by the agent.
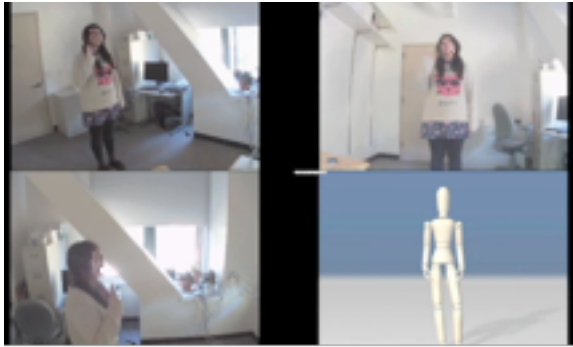
Figure 4: Screenshot of composite feed of video recording.

3. Corpus Characteristics

The corpus includes personality profiles, retellings transcripts, and gesture coding. It also includes the original agent scripts along with gesture coding of the agent's gestures in the same format as the participants.

3.1 Personality Profiles

Information related to the personality profile of the participants is included in the corpus. The short questionnaire the participants filled out provided a measure of personality along the Big 5 personality dimensions through the Ten Item Personality Measure (TIPI; Gosling, Rentfrow, & Swann, Jr., 2003). The participants who participated in this study represented a wide range of personality profiles, including on the dimension of extraversion most relevant to the current analysis.

3.2 Retellings

The story retellings were transcribed for both words and gestures. On average, participants spoke 335 words (SD = 52) across both stories. The verbal transcriptions can be compared to the original agent script for measures of retelling accuracy using an automatic script that produced counts of exact matches. On average, participants introduced 71 words per story that were not found in the original script, while an average of 34 words from the original script were omitted in the retelling. The two original agent scripts were 225 and 257 words long.

3.3 Gesture Coding

Fourteen participants did not gesture at all during the course of the experiment. For the participants who did gesture, the degree of gesturing was highly variable. On average these participants gestured 36 times (SD = 36 gestures), with the number of gestures across the whole experiment ranging from a single gesture to 150

gestures. Transcription and coding were combined in a modified XML format (see Figure 3 below).

3.4 Gesture Analyses

The gesture corpus permits analysis of both the gestural behavior of people interacting with a virtual agent as well as the influence of the agent's gestures on this communicative behavior. It demonstrates the high degree of variability in the expressive behaviors of humans interacting with virtual agents, from those who did not gesture at all to those who displayed a large degree of communicative gesturing. In the following we present an analysis of the participants' gesture style in relation to the degree of extraversion displayed in the agent's gesturing style over time that also works to exemplify how this corpus may be used for future work. As participants did not produce gestures at every level of agent expressive behavior, and exhibited large variation in the degree of gestures produced, a mixed effects model was employed with random intercepts for participant (Snijders & Bosker, 1999). Chi-square tests were run comparing the goodness of fit of models including the variable of interest with null models.

Each of the 1127 total gestures across the corpus was coded for type (see Table 2), handedness, spatial extent, and temporal structure (whether the gesture contained a preparatory phase, hold phase, or retract phase, and the length of these and the gesture stroke phases). Gestures were also coded as to whether they mimicked the gesturing behavior of the agent. Of the 1127 gestures, about a quarter, 285, matched the location of one of the agent's gestures. Of these gestures produced in the same location as the agent's gestures, 131 also matched the form of the agent's gesture.

Table 2: Gesture Type Counts	
Gesture Type	Count
Beat	210
Metaphoric	643
Iconic	211
Deictic	54

A single spatial extent score was calculated for each gesture by averaging across the three dimensions coded for the beginning and end of the gesture. Participants' extraversion level predicted the spatial extent of their gestures, $\chi^2(1) = 10.69$, $p < .01$. Participants who reported higher extraversion produced larger gestures over the course of the total experiment, $\beta=.12$, $SDE = .04$.

```
1    <participant id="25">
2  ▼ <personality ex="4" op="5" ag="5.5" ems="4.5" co="5">
3  ⌐ </personality>
4
5    <story id="0" order="0" version=0>
6  ▼ <part number="1" agentexlvl="1">
7    yeah um so when you were 17 it was a sunday you went over to your boyfriends house and,
...  you got into a really really big fight like an all out <gesture type="metaphoric"
...  handedness="both" prep=".54" stroke=".44" se_onset_height="1" se_onset_width="1"
...  se_onset_radial="3" se_offset_height="1" se_offset_width="1" se_offset_radial="3"
...  retract=".37" agent_match_loc="false" agent_match_form="false"> screaming </gesture> fight
8  ⌐ </part>
9
10 ▼ <part number="2" agentexlvl="2">
11   yeah so, you decided to just leave and you were crying and stuff, um and you got in your
...  car but you are a really good <gesture type="beat" handedness="both" stroke=".78"
...  se_onset_height="1" se_onset_width="1" se_onset_radial="3" se_offset_height="1"
...  se_offset_width="1" se_offset_radial="3" retract=".28" agent_match_loc="false"
...  agent_match_form="false"> driver so </gesture> you don't know why this <gesture
...  type="beat" handedness="both" stroke=".93" se_onset_height="1" se_onset_width="1"
...  se_onset_radial="3" se_offset_height="1" se_offset_width="1" se_offset_radial="3"
...  retract=".12" agent_match_loc="false" agent_match_form="false"> happened to </gesture> you
12 ⌐ </part>
```

Figure 5: Sample presentation of transcription data, providing participant personality profile, as well as verbal and gestural transcription information.

To analyze accommodation to the agent's gestures, we ran separate analyses for the part-by-part retellings and the full story retelling that happened at the end of the experiment. For the part-by-part retellings, we used the agent's extraversion level for each individual part as a predictor of the participant's spatial extent. For the full retelling we measured participants' gesturing style in regards to the two general conditions, which were whether the agent's behavior moved from extraverted to introverted or from introverted to extraverted over the course of the experiment.

For the part-by-part retelling, the agent's extraversion as expressed through their gesture style within a given story part did not significantly impact the participants' spatial extent beyond their own extraversion level, $\chi^2(1) = 0.91$, $p = .34$, nor did these two factors interact, $\chi^2(1) = 2.46$, $p = .29$. We also tested whether participants' behavior varied systematically over time, such as whether they moved more as the experiment went on, and found that this factor did not significantly improve model fit, $\chi^2(1) = 0.54$, $p = .46$. Looking at the likelihood that a participant would repeat a gesture that the agent displayed, or gesture at the same point in the talk at which the agent gestured, neither the agent's extraversion level nor the participants' own reported extraversion significantly improved model fit for either variable, all ps > .05.

Within the full retellings, the agent's general pattern of change in expressive behavior interacted with the participant's extraversion, such that high extraverts and low extraverts displayed distinct patterns of changes in spatial extent across the two stories (one story for each of the two agent expressive behavior patterns). A model including the interaction predicted significantly more variance, $\chi^2(1) = 6.66$, $p < .01$. Participants with high levels of extraversion were more influenced by the agent condition than participants with low extraversion. However, as visible in Figure 4, these high extraverts diverged from the pattern of expressive behavior of the agents, making gestures with larger spatial extent when the agent moved to a more introverted style of nonverbal expressive behavior over the course of the interaction. While the majority of the literature on interactional effects in expressive behavior have focused on convergence (Pickering & Garrod, 2004; Mol et al., 2012), divergence in style is not unprecedented (Giles, Coupland, & Coupland, 1991) and may be an important aspect of interpersonal interaction (Giles & Ogay, 2007).
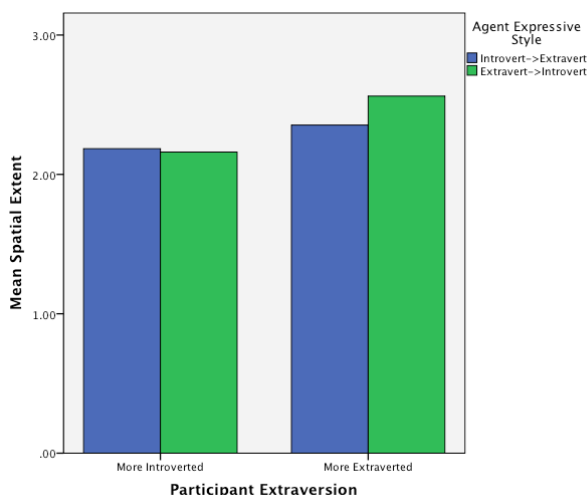

Figure 6: The mean spatial extent of gestures produced by participants during the full retelling of the stories learned from the agent. For ease of visual display, we present low and high extraverts with a median split (statistical tests were on continuous data). The agent's expressive style across the story learning interaction interacted with the participants' extraversion such that high extraverts displayed divergent expressive styles.

4. Conclusion

We have described an experiment to collect a corpus of

people's speech and gestures produced in a spontaneous interaction with a virtual agent. The agent was designed to vary its nonverbal expressive behavior. This corpus captures both the speech and the gestural behaviors of participants learning and retelling a story back to the agent. Participants produced gestures that correlated with their personality profile: The more extraverted they were, the greater the spatial extent of their gestures. In addition, the personality of the participant interacted with the personality correlates visible in the agent's gesturing: Some people produced distinct gesture styles when interacting with agents whose behavior changed over the course of the conversation. Interacting with autonomous conversational agents, in particular those designed to produce nonverbal communicative and expressive behaviors, is a newly developing area. We consider it important to provide data on the human users as well, including the variability in communicative behavior directed towards these agents.

5. Acknowledgements

6. References

André, E., Aylett, R., Hofstede, G. J., & Paiva, A. (2014). Computational Models of Cultural Behavior for Human-Agent Interaction (Dagstuhl Seminar 14131). *Dagstuhl Reports, 4*(3).

André, E., & Rist, T. (2002). From adaptive hypertext to personalized web companions. *Communications of the ACM, 45*(5), 43-46.

Aronson, E., & Linder, D. (1965). Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology, 1,* 156-171.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16,* 814-819.

Bailenson, J. N., Yee, N., Patel, K., & Beall, A. C. (2008). Detecting digital chameleons. *Computers in Human Behavior, 24,* 66-87.

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58,* 495-520.

Bergmann, K. & S. Kopp (2012). Gestural Alignment in Natural Dialogue. In D. Peebles, N. Miyake & R.P. Cooper (Eds.), *Proc. of the 34th Annual Conf. of the Cognitive Science Society (CogSci 2012* p. 1326–1331.

Bickmore, T. (2002). Towards the design of multimodal interfaces for handheld conversational characters. *CHI 2002.*

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Nass, C. I. (2003). Syntactic alignment between computers and people: The role of belief about mental states. *Proc. of the 25th Annual Meeting of the Cognitive Science Society,* 186-191.

Brugman, H., & Russel, A. (2004). Annotating Multimedia/Mutli-modal resources with ELAN. In: *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation.*

Buschmeier, H., Bergmann, K., & Kopp, S. (2009). An alignment-capable microplanner for natural language generation. In: *Proc. of the 12th European Workshop on Natural Language Generation,* (pp. 82-89).

Chartranand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893-910.

De Jong, M., Theune, M., & Hofs, D. (2008). Politeness and alignment in dialogues with a virtual agent. In: *Proc. of the Seventh Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008),* 207-214.

Endrass, Birgit, Ionut Damian, Peter Huber, Matthias Rehm, and Elisabeth André. 2010.) "Generating culture-specific gestures for virtual agent dialogs." In *Intelligent Virtual Agents*, pp. 329-335.

Funder, D.C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 53,* 409-418.

Giles, H., Couplan, N., & Coupland, J. (1991). Communication, context, and consequence. In: H. Giles, N. Coupland, & J. Coupland (Eds.), *Contexts of accommodation* (pp. 49-68). New York: Cambridge University Press.

Giles, H., & Ogay, T. (2007). Communication Accommodation Theory. In: B. Whaley & W. Samter (Eds.), *Explaining communication: Contemporary theories and exemplars.* Mahwah, NJ: Lawrence Erlbaum.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big Five Personality domains. *Journal of Research in Personality, 37,* 504-528.

Hartmann, B., Mancini, M., & Pelachaud, C. (2005). Implementing expressive gesture synthesis for embodied conversational agents. In *gesture in human-Computer Interaction and Simulation* (pp. 188-199). Springer Berlin Heidelberg.

Heyselaar, E., Hagoort, P., & Segaert, K. (2014). In dialogue with an avatar, syntax production is identical compared to dialogue with a human partner. In: *Proceedings of the 36th Annual Meeting of the Cognitive Science Society,* 2351-2356.

Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of nonverbal behavior, 35,* 133-153.

Hu, C., Walker, M. A., Neff, M., & Fox Tree, J. E. (2015). Storytelling agents with personality and adaptivity. In: *Intelligent Virtual Agents* (pp. 181-193). Springer International Publishing.

Hu, Z., Dick, M., Chang, C.N., Bowden, K., Neff, M., Fox Tree, J. E., and Walker, M. A. (2016). A corpus of gesture annotated dialogues for monologue

to dialogue generation from personal narratives. *Language Resources and Evaluation Conf.* (LREC).

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61,* 521-551.

Kimbara, I. (2008). Gesture form convergence in joint description. *Journal of Nonverbal Behavior, 32,* 123-131.

Kipp, M., Neff, M., & Albrecht, I. (2006). An annotation scheme for conversational gestures: How to economically capture timing and form. In: *Proceedings of the Workshop on "Multimodal Corpora" at LREC,* 24-27.

Kipp, M., Neff, M., & Albrecht, I. (2007). An annotation scheme for conversational gestures: how to economically capture timing and form. *Language Resources and Evaluation, 41*(3-4), 325-339

Kipp, M., Neff, M., Kipp, K. H., & Albrecht, I. (2007) Towards natural gesture synthesis: Evaluating gesture units in a data-driven approach to gesture synthesis. *Intelligent Virtual Agents* (pp. 15-28).

Kita, S., van Gijn, I., & van der Hulst, H. (21998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In: Wachsmuth, I., Fröhlich, M. (Eds.) *Gesture and Sign Language in Human-Computer Interaction. LNCS (LNAI), 1371.*

Kopp, S., B. Krenn, S. Marsella, A. N. Marshall, C. Pelachaud, H. Pirker, K. R. Thórisson, and H. Vilhjálmsson. Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent virtual agents*, pp. 205-217.

Kramer, N., Simons, N., & Kopp, S. (2007). The effects of an embodied conversational agent's nonverbal behavior on user's evaluation and behavioral mimicry. *Proc. of Intelligent Virtual Agents 2007* (pp. 238-251).

LaFrance, B. H., Heisel, A. D., & Beatty, M. J. (2004). Is there empirical evidence for a nonverbal profile of extraversion?: A meta-analysis and critique of the literature. *Communication Monographs, 71,* 38-48.

Liu, K., Tolins, J., Fox Tree, J. E., Walker, M., & Neff, M. (2013). Judging IVA personality using an open-ended question. In *Lecture Notes in Artificial Intelligence* LNAI 8108 (pp. 396–405).

Liu, K., Tolins, J., Fox Tree, J. E., Neff, M., & Walker, M. A. (2016). Two techniques for assessing virtual agent personality. *IEEE Transactions on Affective Computing, 7*(1), 94-105.

Lippa, R. (1998). The nonverbal display and judgment of extraversion, masculinity, femininity, and gender diagnosticity: A lens model analysis. *Journal of Research in Personality, 32,* 80-107.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication in synchronized. *Cognitive Science, 36* 1-24.

Mairesse, F., & Walker, M. A. (2010). Towards personality-based user adaptation: psychologically informed stylistic language generation. *User Modeling and User-Adapted Interaction*, *20*(3), 227-278.

Mairesse, F., & Walker, M. A. (2011). Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics*,*37*(3), 455-488.

Stephanie M. Lukin, Kevin Bowden, Casey Barackman, and Marilyn A. Walker. 2016. PersonaBank: A Corpus of Personal Narratives and Their Story Intention Graphs. *LREC 2016.*

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: The University of Chicago Press.

McNeill, D. (2005). *Gesture and thought.* Chicago: University of Chicago Press.

Mol, L, Krahmer, E., Maes, A., & Swerts, M. (2012). Adaption in gesture: Converging hands or converging minds? *Journal of Memory and Language, 66,* 249-264.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56,* 81-103.

Neff, M., Wang, Y., Abbott, R., & Walker, M. (2010). Evaluating the effect of gesture and language on personality perception in conversational agents. *Proceedings of Intelligent Virtual Agents (IVA '10).*

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27,* 19-190.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7,* 226-231.

Rieser, V., & Lemon, O. (2010). Natural language generation as planning under uncertainty for spoken dialogue systems. In *Empirical methods in natural language generation* (pp. 105-120).

Riggio, R. E., & Friedman, H. S. (1986). Impression formation: The role of expressive behavior. *Journal of Personality and Social Psycology, 50,* 421-427.

Salem, M., Kopp, S., Wachsmuth, I., Rohlfing, K., & Joublin, F. (2012). Generation and evaluation of communicative robot gesture. *International Journal of Social Robotics, 4*(2), 201-217.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis.* London: Sage Publications.

Tolins, J., Liu, K., Wang, Y., Fox Tree, J. E., Walker, M. Neff, M. (2016). A multimodal corpus of matched and mismatched extravert-introvert conversational pairs. In: *Proceedings of the Language Resources and Evaluation Conference (LREC).*

Tolins, J., Liu, K., Wang, Y., Fox Tree, J. E., Walker, M. Neff, M. (2013). Gestural adaptation in extravert-introvert pairs and implications for IVAs. In R. Aylett et al. (Eds.), *Lecture Notes in Artificial Intelligence,* LNAI 8108 (pp. 481–482).

Walker, M. A., Stent, A., Mairesse, F., & Prasad, R. (2007). Individual and domain adaptation in sentence planning for dialogue. *Journal of Artificial Intelligence Research*, 413-456.