A Corpus of Gesture-Annotated Dialogues for Monologue-to-Dialogue Generation from Personal Narratives

Zhichao Hu[1], Michelle Dick[1], Chung-Ning Chang[1], Kevin Bowden[1]
Michael Neff[2], Jean E. Fox Tree[1], Marilyn Walker[1]
University of California, Santa Cruz[1]; University of California, Davis[2]
1156 High Street, Santa Cruz, CA, USA; 1 Shields Ave, Davis, CA, USA
{zhu, mdick, cchan60, kkbowden}@ucsc.edu, mpneff@ucdavis.edu, {foxtree, mawalker}@ucsc.edu

Abstract

Story-telling is a fundamental and prevalent aspect of human social behavior. In the wild, stories are told conversationally in social settings, often as a dialogue and with accompanying gestures and other nonverbal behavior. This paper presents a new corpus, the STORY DIALOGUE WITH GESTURES (SDG) corpus, consisting of 50 personal narratives regenerated as dialogues, complete with annotations of gesture placement and accompanying gesture forms. The corpus includes dialogues generated by human annotators, gesture annotations on the human generated dialogues, videos of story dialogues generated from this representation, video clips of each gesture used in the gesture annotations, and annotations of the original personal narratives with a deep representation of story called a STORY INTENTION GRAPH. Our long term goal is the automatic generation of story co-tellings as animated dialogues from the STORY INTENTION GRAPH. We expect this corpus to be a useful resource for researchers interested in natural language generation, intelligent virtual agents, generation of nonverbal behavior, and story and narrative representations.

Keywords: storytelling, personal narrative, dialogue, virtual agents, gesture generation, personality

1. Introduction

Sharing experiences by story-telling is a fundamental and prevalent aspect of human social behavior (Bruner, 1991; Bohanek et al., 2006; Labov and Waletzky, 1997; Nelson, 2003). In the wild, stories are told conversationally in social settings, often as a dialogue and with accompanying gestures and other nonverbal behavior (Tolins et al., 2016a). Story-telling in the wild serves many different social functions: e.g. stories are used to persuade, share troubles, establish shared values, learn social behaviors, and entertain (Ryokai et al., 2003; Pennebaker and Seagal, 1999; Gratch et al., 2013).

Previous research has also shown that conveying information in the form of a dialogue is more engaging, effective and persuasive compared to a monologue (Lee et al., 1998; Craig et al., 2000; Suzuki and Yamada, 2004; André et al., 2000). Thus our long term goal is the automatic generation of story co-tellings as animated dialogues. Given a deep representation of the story, many different versions of the story can be generated, both dialogic and monologic (Lukin and Walker, 2015; Lukin et al., 2015).

This paper presents a new corpus, the STORY DIALOGUE WITH GESTURES (SDG) corpus, consisting of 50 personal narratives regenerated as dialogues by human annotators, complete with annotations of gesture placement and accompanying gesture forms. These annotations can be supplemented programmatically to produce changes in the nonverbal dialogic behavior in gesture rate, expanse and speed. We have thus used these to generate tellings that vary the personality of the teller (introverted vs. extroverted) (Hu et al., 2015). An example original monologic personal narrative about having cats as pets is shown in Figure 1. The human-generated dialogue corresponding to Figure 1 is shown in Figure 2.

Pet Story
I have two cats. I always felt like I was a dog person, but I decided to get a kitty because they are more low maintenance than dogs. I went to a no-kill shelter to get our first cat. I wanted a little kitty, but the only baby kitten they had scratched the crap out of me the minute I picked it up. SO, that was a big “NO”. They had what they called “teenagers”. They were cats that were 4-6 months old. Not adults, but a little bigger than the little kittens. One stood out - mostly because she jumped up on a shelf behind my husband and smacked him in the head with her paw. I had a winner! I had no idea how much personality a cat can have. Our first kitty loves to play. She will play until she is out of breath. Then, she looks at you as if to say, “Just give me a minute, I’ll get my breath back and be good to go.” Sometimes I wish I had that much enthusiasm for anything in my life. She loves to chase a string. It’s the best thing ever. Ok, maybe it runs a close second to hair scrunchies. I play fetch with my hair scrunchies. I throw them down the stairs and she runs (top speed) to get them and bring them back. Again, she will do this until she is out of breath. If only I could work out that hard... I’d probably be thinner.

Figure 1: An example story on the topic of Pet.

The SDG corpus includes 50 dialogues generated by human annotators, gesture annotations on human generated dialogues, videos of story dialogues generated from this representation that vary the introversion and extraversion of the animated agents, video clips of each gesture used in the gesture annotations, and annotations of the original personal narratives with a deep representation of story called a STORY INTENTION GRAPH or SIG (Elson, 2012a; Elson

Pet Story Dialogue	
A1:	I have always felt like I was a dog person but our two cats are great. They are much more low maintenance than dogs are.
B1:	Yeah, I'm really glad we got our first one at a no-kill shelter.
A2:	I had wanted a little kitty, but the only baby kitten they had scratched the crap out of me the minute I picked it up so that was a big "NO".
B2:	Well, the no-kill shelter also had what they called "teenagers", which were cats around four to six months old... A bit bigger than the little kitties.
A3:	Oh yeah, I saw those "teenagers". They weren't exactly adults, but they were a bit bigger than the little kittens.
B3:	Yeah one of them really stood out to me then - mostly because she jumped up on a shelf behind us and smacked me in the head with her paw.
A4:	Yeah, we definitely had a winner!
B4:	I had no idea how much personality a cat can have. Our first kitty loves playing. She will play until she is out of breath.
A5:	Yeah, and then after playing for a long time she likes to look at you like she's saying, "Just give me a minute, I'll get my breath back and be good to go."
B5:	Sometimes I wish I had that much enthusiasm for anything in my life.
A6:	Yeah, me too. Man, she has so much enthusiasm for chasing string too! To her it's the best thing ever. Well ok, maybe it runs a close second to hair scrunchies.
B6:	Oh I love playing fetch with her with hair scrunchies!
A7:	Yeah, you can just throw the scrunchies down the stairs and she runs at top speed to fetch them. And she always does this until she's out of breath!
B7:	If only I could work out that hard before I was out of breath... I'd probably be thinner.

Figure 2: Manually constructed dialogue from the pet story in Figure 1.

and McKeown, 2010; Elson, 2012b). We expect this corpus to be a useful resource for researchers interested in natural language generation, intelligent virtual agents, generation of nonverbal behavior, and story and narrative representations. Section 2. provides an overview of the original corpus of monologic personal narrative blog posts. Section 3. describes how the human annotators generated dialogues from the personal narratives. Section 4. describes the gesture annotation process and provides more details of the gesture library. We conclude by summarizing the paper and discussing possible applications of the corpus in Section 6..

2. Personal Narrative Monologic Corpus

We have selected 50 personal stories from a subset of personal narratives extracted from the corpus of blogs included in the ICWSM 2010 dataset challenge (Kevin Burton and Soboroff, 2009; Gordon and Swanson, 2009). We manually selected stories that are suitable for retelling as dialogues, i.e. where the events that are discussed in the stories could have been experienced by more than one person. The story topics include camping, holiday, gardening, storms, and parties. Table 1 shows the number of stories in each topic. Each story ranges from 174 words to 410 words. A sample story about pets is shown in Figure 1 and another story about being present at a protest is shown in Figure 5.

Topic	Number of Stories
Camping	3
Holiday	5
Gardening	7
Party	10
Pet	3
Sports	4
Travel	7
Weather Conditions	3
Other	13

Table 1: Distribution of story topics in the SDG corpus.

Although we are not making using of the STORY INTENTION GRAPHS (SIGs) yet to automatically produce dialogues from their monologic representations, we have annotated each of the 50 stories with their SIG using the freely available annotation tool Scheherazade (Elson and McKeown, 2009). More description of the Scheherazade annotation and the resulting SIG representation is provided in our companion paper (Lukin et al., 2016). Our approach builds on the DramaBank language resource, a collection classic stories that also utilize the SIG representation (Elson, 2012a; Elson and McKeown, 2010; Elson, 2012b), but the SIG formalism has not previously been used for the purpose of automatically generating animated dialogues, and this is one of the first uses of the SIG on personal narratives (Lukin and Walker, 2015; Lukin et al., 2015).

DramaBank provides a symbolic annotation tool for stories called Scheherazade that automatically produces the SIG as a result of the annotation. Every annotation involves: (1) identifying key entities that function as characters and props in the story; and (2) modeling events and statives as propositions and arranging them in a timeline. Currently our Scheherazade annotations only contain the timeline layers.

3. Dialogue Annotations

In the wild, stories are told conversationally in social settings, and in general research has shown that conveying information in the form of a dialogue is more engaging and memorable (Lee et al., 1998; Craig et al., 2000; Suzuki and Yamada, 2004; André et al., 2000). In story-telling and at least some educational settings, dialogues have cognitive advantages over monologues for learning and memory. Students learn better from a verbally interactive agent than from reading text, and they also learned better when they interacted with the agent with a personalized dialogue (whether spoken or written) than a non-personalized monologue (Moreno et al., 2000).

Our goal with the dialogue annotations is to generate a natural dialogue from the original monologic text, with the long

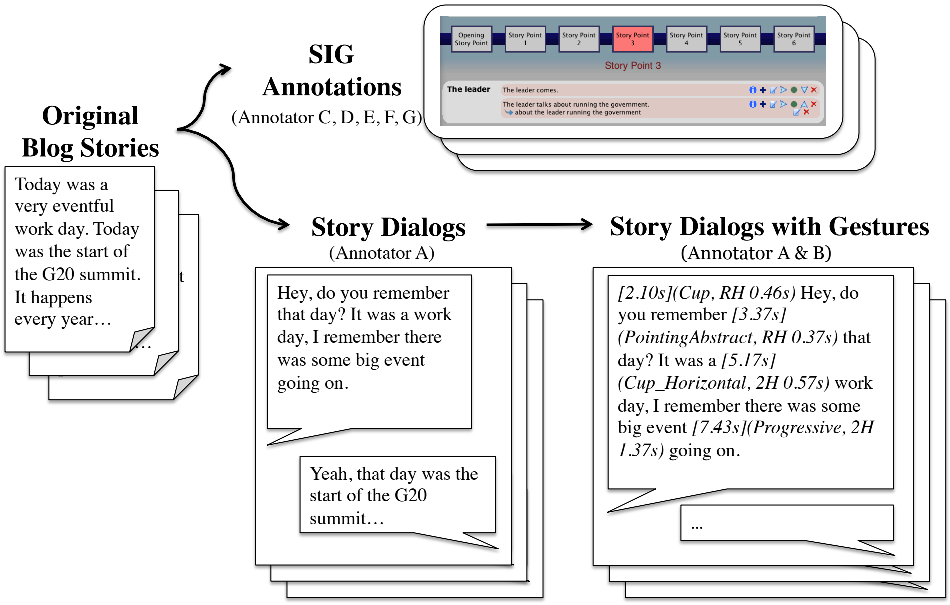

Figure 3: Overview of the SDG corpus.

Protest Story SIG Annotations

Story Elements:
Characters: the government, the people, the leader, the police, the society, the street, the summit meeting (name: G20 Summit)
Props: the bullet, the police car, the tear gas, the thing, the view, the window
Qualities: peace, prosperity, riot, today
Behaviors: the leader running the government, the people protesting

Timeline:

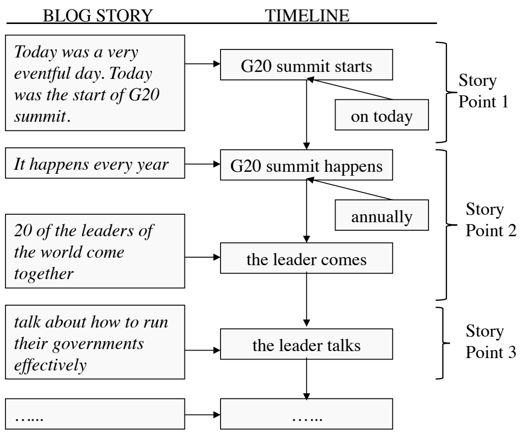

Figure 4: Part of Scheherazade annotations for the protest story in Figure 5.

Protest Story

Today was a very eventful work day. Today was the start of the G20 summit. It happens every year and it is where 20 of the leaders of the world come together to talk about how to run their governments effectively and what not. Since there are so many leaders coming together their are going to be a lot of people who have different views on how to run the government they follow so they protest. There was a protest that happened along the street where I work and at first it looked peaceful until a bunch of people started rebelling and creating a riot. Police cars were burned and things were thrown at cops. Police were in full riot gear to alleviate the violence. As things got worse tear gas and bean bag bullets were fired at the rioters while they smash windows of stores. And this all happened right in front of my store which was kind of scary but it was kind of interesting since I've never seen a riot before.

Figure 5: An example story on the topic of Protest.

term aim of using these human-generated dialogues to guide the development of an automatic monologue-to-dialogue generation engine. First, one trained annotator writes all the stories into two-person dialogues as illustrated in Figure 2 and Figure 6. The goal of the annotation process is to create natural dialogues by: (1) adding oral language such as acknowledgements and discourse markers, and breaking longer sentences into shorter ones; (2) adding repetitions and confirmations between speakers, which are common in human dialogue, and can also be used as locations for inserting gesture entrainment; (3) re-using phrases in the original story, but changing or deleting content that doesn't fit the storytelling setting; (4) making the story sound like the two speakers experience the event together.

The audio for each of the agents is then produced by running the human-generated dialogue turns through the AT&T Text to Speech engine, one turn at a time. Speaker A uses AT&T's female voice Crystal, Speaker B uses AT&T's male voice Mike. At the beginning of the audio for each story, a two-

Protest Story Dialogue	
A1:	Hey, do you remember that day? It was a work day, I remember there was some big event going on.
B1:	Yeah, that day was the start of the G20 summit. It's an event that happens every year.
A2:	Oh yeah, right, it's that meeting where 20 of the leaders of the world come together. They talk about how to run their governments effectively.
B2:	Yeah, exactly. There were many leaders coming together. They had some pretty different ideas about what's the best way to run a government.
A3:	And the people who follow the governments also have different ideas. Whenever world leaders meet, there will be protesters expressing different opinions. I remember the protest that happened just along the street where we work.
B3:	It looked peaceful at the beginning....
A4:	Right, until a bunch of people started rebelling and creating a riot.
B4:	Oh my gosh, it was such a riot, police cars were burned, and things were thrown at cops.
A5:	Police were in full riot gear to stop the violence.
B5:	Yeah, they were. When things got worse, the protesters smashed the windows of stores.
A6:	Uh huh. And then police fired tear gas and bean bag bullets.
B6:	That's right, tear gas and bean bag bullets... It all happened right in front of our store.
A7:	That's so scary.
B7:	It was kind of scary, but I had never seen a riot before, so it was kind of interesting for me.

Figure 6: Manually constructed dialogue from the protest story in Figure 5.

second blank audio is inserted in order to give the audience time to prepare to listen to the story before it starts. The timeline for the audio is then used to annotate the beginning timestamps for the gestures: as shown in Figures 9 and 10, in front of every gesture, a time is shown inside a pair of square brackets to indicate the beginning time of this gesture stroke. Thus any change in the wording of the dialogic telling requires regenerating the audio and relabelling the gesture placement. In our envisioned future Monologue-to-Dialogue generator, both gesture placement and gesture form would be automatically determined.

4. Gesture Annotations

We annotate the dialogues with a gesture tag that specifies the gesture form (e.g. a pointing gesture or a conduit gesture). We also specify gesture start times, but do not specify stylistic variations that can be applied to particular gestures (e.g. gesture expanse, height and speed). Each story has two versions of annotations done by different annotators. The annotators are advised to insert a gesture when the dialogue introduces new concepts, and add gesture adaptation (mimicry) when there are repetitions or confirmations in the dialogue. The decisions of where to insert a gesture and which gesture to insert are mainly subjective.

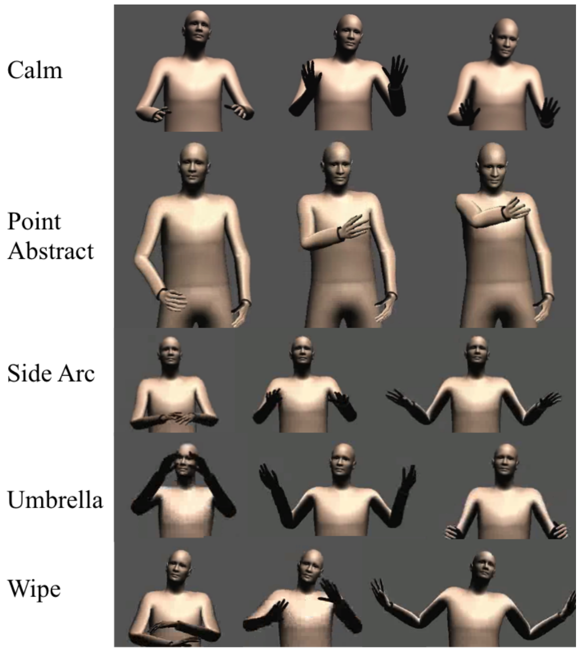

Figure 7: A subset of the 271 gestures in our gesture library that can be used in annotation and produced in the animations.

We use gestures from a database of 271 motion captured gestures, including metaphoric, iconic, deictic and beat gestures. The videos of these gestures are included in the corpus. Gesture capture subjects were all native English speakers. But given a database of similarly categorized gestures from a different culture, our corpus could be used to generate culture specific performances of these stories (Neff et al., 2008; Rehm et al., 2009).
Figure 7 provides samples of the range of gestures in the gesture database, and Figure 8 illustrates how every gesture can be generated to include up to 4 phases (Kita et al., 1998; Kendon, 1980):

- prep: move arms from default resting position or the end point of the last gesture to the start position of the stroke
- stroke: perform the movement that conveys most of the gesture's meaning
- hold: remain at the final position in the stroke
- retract: move arms from the previous position to a default resting position

Figure 9 shows the first 5 turns of the protest story annotated with gestures from the library. Figure 10 shows the first 6 turns of the pet story annotated with gestures. The timing information of the gestures comes from the TTS audio timeline. Each gesture annotation contains information in the following format: ([gesture stroke begin time]gesture name, hand use[stroke duration]). For example, in the first gesture "([1.90s]Cup, RH [0.46s])", gesture stroke begins at 1.9 seconds of the dialogue audio, it is a "Cup" gesture, uses the right hand, and the gesture stroke lasts 0.46 seconds.

Figure 8: Prep, stroke, hold and retract phases of gesture "Cup_Horizontal".

Protest Story w/Gestures

A1: ([1.90s]Cup, RH[0.46s]) Hey, do you remember ([3.17s]PointingAbstract, RH[0.37s]) that day? It was a ([4.97s]Cup_Horizontal, 2H[0.57s]) work day, I remember there was some big event ([7.23s]SweepSide1, RH[0.35s]) going on.

B1: Yeah, that day was the start of ([9.43s]Cup_Down_alt, 2H[0.21s]) the G20 summit. It's an event that happens ([12.55s]CupBeats_Small, 2H[0.37s]) every year.

A2: Oh yeah, ([14.2s]Cup_Vert, RH[0.54s]) right, it's that meeting where 20 of the leaders of the world ([17.31s]**Regressive, RH[1.14s]**) come together. They talk about how to run their governments ([20.72s]**Cup, RH[0.46s]**) effectively.

B2: Yeah, ([22.08s]**Cup_Up, 2H[0.34s]**) exactly. There were many leaders ([24.38s]**Regressive, LH[1.14s] /** Eruptive, LH[0.76s]) coming together. They had some pretty ([26.77s]**WeighOptions, 2H[0.6s]**) different ideas about what's the best way to *([29.13s]**Cup, RH[0.46s] /** ShortProgressive, RH[0.38s]) run a government.

A3: And *([30.25s]PointingAbstract, RH[0.37s]) the people who follow the governments also have ([32.56s]**WeighOptions, 2H[0.6s] /** Cup, 2H[0.46s]) different ideas. Whenever ([34.67s]**Cup_Up, 2H[0.34s] /** Dismiss, 2H[0.47s]) world leaders meet, there will be protesters expressing ([37.80s]Away, 2H[0.4s]) different opinions. I remember the *([39.87s]Reject, RH[0.44s]) protest that happened just ([41.28s]SideArc, 2H[0.57s]) along the street where we work.

B3: ......

Figure 9: Protest dialogue with gesture annotations. Pictures show the first 6 gestures in the dialogue.

Research has shown that people prefer gestures occurring earlier than the accompanying speech (Wang and Neff, 2013). Thus in this annotation, a gesture stroke is positioned 0.2 seconds before the beginning of the gesture's following word. For example, the first word after gesture "Cup" is "Hey", it begins at 2.1 seconds, then the stroke of gesture "Cup" begins at 1.9 seconds. Each story dialogue has two versions of gesture annotations from different annotators.

Our gesture annotation does not specify stylistic variations that should be applied to particular gestures. We used custom animation software that can vary the amplitude, direction and speed in order to affect stylistic change. The default gesture annotation frequency is designed for extraverts, with a gesture rate of 1 - 3 gestures per sentence. For an introverted agent, a lower gesture rate can be achieved by removing some of the gestures. In this way, both speakers' gestural performance can vary from introverted to extraverted using the whole scale of parameter values for every parameter.

In addition, we can also vary gestural adaptation in the annotation. For example, in extravert & extravert gestural adaptation (based on the model and data described in Tolins et al. (2013), Tolins et al. (2016b) and Neff et al. (2010)), two extraverts move together towards a more extraverted personality. Gesture rate is increased by adding extra gestures (marked with an asterisk "*"). Specific gestures are copied as part of adaptation, especially when the co-telling involves repetition and confirmation. Gestures in bold indicate copying of gesture form (adaptation), gestures after the

Pet Story w/Gestures

A1: I ([2.31s]Dismiss1_AntonyTired, RH[0.54s])have always felt like I was a dog person but our two cats are great. They are ([7.50s]Cup5_ShakesTired2_FingerSkel, RH[0.23s]) much more low maintenance than dogs are.
B1: Yeah, I'm really glad we got ([12.44s]SideOut1, 2H[0.29s]) our first one at a no-kill shelter.
A2: I had wanted a little kitty, ([16.44s]SweepSide1, RH[0.35s]) but the only baby kitten they had ([17.90s]BackHandBeats_Tight, 2H[0.37s]) scratched the crap out of me the minute I ([19.75s]Shovel, 2H[0.17s]) picked it up so that was a big "NO".
B2: Well, ([22.69s]Cup_Down_alt, RH[0.21s]) the no-kill shelter also had what they called ([24.70s]**Cup_Horizontal , RH[0.57s]**) "teenagers", which were cats around four to six months old... a bit bigger than the ([28.79s]**ShyCalmShake, 2H [0.56s]**) little kitties.
A3: Oh yeah, I saw those ([31.24s]**Cup_Horizontal, RH[0.57s]** / SideOut_vibrate, 2H[0.33s]) "teenagers". They *([32.75s]HandToChest_Vibrate, 2H[0.41s]) weren't exactly adults, but they were a bit ([35.78s]**ShyCalmShake, 2H [0.56s]** /SideOut1, LH[0.29s]) bigger than the little kittens.
B3: Yeah ([36.42s]SideOut_vibrate, 2H[0.33s]/ PointingHere, RH[0.36]) one of them really stood out to me then - *([38.92s]Cup11_ShakesTired2_FingerSkel, RH[0.30s]) mostly because she *([40.02s]CupBeats_Small, 2H[0.37s]) jumped up on a shelf behind us and ([42.20s]SideOut1, LH[0.29s]/ GraspSmall, RH[0.18s]) smacked me in the head with her paw.
A4: ......

Figure 10: Pet dialogue with gesture annotations. Pictures show the first 6 gestures in the dialogue.

slash "/" are non-adapted.
Combined with personality variations for gestures described in the previous paragraph, it is possible to produce combinations of two agents with any level of extraversion engaged in a conversation with or without gestural adaptation.

5. Possible Applications

With various types of annotations, the SDG corpus is useful in many applications. We introduce two possible areas of applications: conducting gestural perception experiments and helping with monologue to dialogue generation.

5.1. Gestural Perception Experiments

Our own work to date has used the human-generated dialogues and the gesture annotations to carry out two experiments (Hu et al., 2015): a personality experiment aiming to elicit subjects' perceptions of two virtual agents designed to have different personalities, and an adaptation experiment aiming to find out whether subjects prefer adaptive vs. non adaptive agents.
In the personality experiment, we used 4 stories from the SDG corpus shown in Table 2. We prepared two versions of the video of the story co-telling for each of the four stories, one where the female is extraverted (higher values for gesture rate, gesture expanse, height, outwardness, speed and scale) and the male is introverted (lower values for those gesture features) and one where only the genders (virtual agent model and voice) of the agents are switched. We conducted a between-subjects experiment on Mechanical Turk where we ask workers to answer the TIPI (Gosling et al., 2003) personality survey for one of the agents in the video. Table 2 shows the extraversion scores of agents. The results show that subjects clearly perceive the intended extraverted or intended introverted personality of the two agents ($F = 67.1$, $p < .001$), and there is no significant variation by agent gender ($F = 2.3$, $p = .14$).

Story	Introverted Agent	Extraverted Agent
Garden	4.2	5.4
Pet	4.7	5.0
Protest	4.2	5.3
Storm	3.7	5.7

Table 2: Experiment results: participant evaluated extraversion scores (range from 1 - 7, with 1 being the most introverted and 7 being the most extraverted).

In the gestural adaptation experiment, both agents are designed to be extroverted. The adaptation model between two extraverted speakers is based on Tolins et al. (2013) and Neff et al. (2010) (where both agents become more extraverted). We use the same four stories for the personality experiment. The stimuli for one task has two variations: adapted and non-adapted. Both stimuli use the same audio, contain 2 to 4 dialogue turns with the same gestures as an introduction to the story, and the next (and last) dialogue turn with gesture adaptation or without gesture adaptation. Adaptation only begins to occur in the last dialogue turn. In this way, subjects can get to know the story through the context, and compare the responses to decide whether they like the adapted or non-adapted version. Pictures of our experiment stimuli are shown in Figure 9 and 10. Our results in Table 3 show that subjects prefer adaptive stories. We also include in the SDG corpus all the stimuli used in the experiments, as well as the virtual agents' Ten Item Personality Measure (TIPI) scores (Gosling et al., 2003) rated by the participants.

Story	#A	#NA	%A	%NA
Garden	31	11	73.81%	26.19%
Pet	29	18	61.70%	29.30%
Protest	19	22	46.34%	53.66%
Storm	30	9	76.92%	23.08%
Total	109	60	64%	36%

Table 3: Experiment results: number and percentage of subjects who preferred the adapted (A) stimulus and the non-adapted (NA) stimulus.

Using our SDG corpus, experiments similar to the personality experiment and the adaptation experiment can also be conducted with different personality types. For example, previous work has shown that self-adaptors impact the perception of neuroticism (e.g. "rub_forehead", "scratch_face" and "scratch_neck" gestures in our database) (Neff et al., 2011) and gesture disfluency was one of a set of factors that collectively increase neuroticism (Liu et al., 2016). It is not clear how these factors may vary in two person interaction. There are also other dimensions of personality for which there are not well understood movement correlates. Agreeableness, for example, is both not well understood from a movement perspective and a trait that likely benefits from being studied in a two character setting.

Another interesting possibility would be to study different character combinations in this shared story telling task. Characters might vary in personality, as discussed above and also in physical appearance, gender, size - even species. There has been limited study of animation for two character interaction, so this corpus provides a useful set of test scenarios.

5.2. Monologue to Dialogue Generation

Using SIG annotations and the dialogues, this corpus can be used as a gold standard in story monologue to dialogue conversion. Figure 11 shows preliminary results of our Monologue to Dialogue (M2D) generation system. Our system takes SIG annotations and converts them to Deep Syntactic Structures (DSyntS) used by surface realizers such as RealPro(Rishes et al., 2013). To make the dialogue more oral, we add dialogic features such as:

- Hedge insertion: "technically" in A1 and B3, "I mean" in B1
- Synonym replacement using WordNet: "world" to "global" in B1, "create" to "make" in B2, "fire" to "discharge" in A4
- Question answering interaction: A3 and B3

Automatically Generated Garden Story Dialogue	
A1:	G20 summit annually happened and started on the technically eventful today.
B1:	I mean, the world leader came. The global leader talked of the leader ran the government about.
A2:	The people protested because the people disagreed about the view peacefully and on the street.
B2:	The people made the riot, rebelled and created the rioting. The people burned the police car and threw the thing at the police.
A3:	What alleviated the people of the riot?
B3:	Technically, yeah, the police alleviated the people of the riot. The people smashed the window. The police fired the tear gas at the people.
A4:	The police discharged the bullet at the people.

Figure 11: Automatically generated dialogue using the Protest Story SIG Annotations in Figure 4 .

We are still exploring more features and methods to make the dialogue more smooth and natural. To evaluate M2D generation results, we plan to carry out surveys that ask participants to rank variations of generated dialogue excerpts together with human written dialogue excerpts.

6. Discussion and Conclusion

This paper presents the SDG corpus of annotated blog stories, which contains narrative structure annotations, manually written dialogues, and gesture annotations with personality and adaptation variations. With the combination of different annotation types, we hope this resource will be of use to other researchers interested in dialogue, storytelling, language generation and virtual agents.

7. Acknowledgement

This research was supported by NSF CISE RI EAGER #IIS-1044693, NSF CISE CreativeIT #IIS-1002921, NSF IIS 1115872, Nuance Foundation Grant SC-14-74, and auxiliary REU supplements.

8. Bibliographical References

André, E., Rist, T., van Mulken, S., Klesen, M., and Baldes, S. (2000). The automated design of believable dialogues for animated presentation teams. *Embodied conversational agents*, pages 220–255.

Bohanek, J. G., Marin, K. A., Fivush, R., and Duke, M. P. (2006). Family narrative interaction and children's sense of self. *Family process*, 45(1):39–54.

Bruner, J. (1991). The narrative construction of reality. *Critical inquiry*, pages 1–21.

Craig, S., Gholson, B., Ventura, M., and Graesser, A. (2000). the tutoring research group: Overhearing dialogues and monologues in virtual tutoring sessions. *Intl Journal of Artificial Intelligence in Education*, 11:242–253.

Elson, D. K. and McKeown, K. R. (2009). A tool for deep semantic encoding of narrative texts. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 9–12. Association for Computational Linguistics.

Elson, D. and McKeown, K. (2010). Automatic attribution of quoted speech in literary narrative. In *Proceedings of AAAI*.

Elson, D. (2012a). Dramabank: Annotating agency in narrative discourse. In *LREC*, pages 2813–2819.

Elson, D. K. (2012b). Detecting story analogies from annotations of time, action and agency. In *Proceedings of the LREC 2012 Workshop on Computational Models of Narrative, Istanbul, Turkey*.

Gordon, A. and Swanson, R. (2009). Identifying personal stories in millions of weblog entries. In *Third International Conference on Weblogs and Social Media, Data Challenge Workshop, San Jose, CA*.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the big five personality domains. *Journal of Research in Personality*, 37:504–528.

Gratch, J., Morency, L.-P., Scherer, S., Stratou, G., Boberg, J., Koenig, S., Adamson, T., Rizzo, A. A., et al. (2013). User-state sensing for virtual health agents and telehealth applications. In *MMVR*, pages 151–157.

Hu, Z., Walker, M. A., Neff, M., and Fox Tree, J. E. (2015). Storytelling agents with personality and adaptivity. In *Intelligent Virtual Agents*, pages 181–193. Springer.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. *The relationship of verbal and nonverbal communication*, 25:207–227.

Kevin Burton, A. J. and Soboroff, I. (2009). The icwsm 2009 spinn3r dataset.

Kita, S., Van Gijn, I., and Van Der Hulst, H. (1998). Movement phase in signs and co-speech gestures, and their transcriptions by human coders. In *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*, pages 23–35. Springer-Verlag.

Labov, W. and Waletzky, J. (1997). Narrative analysis: Oral versions of personal experience.

Lee, J., Dineen, F., and McKendree, J. (1998). Supporting student discussions: it isn't just talk. *Education and Information Technologies*, 3(3-4):217–229.

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

Lukin, S. M. and Walker, M. A. (2015). Narrative variations in a virtual storyteller. In *Intelligent Virtual Agents*, pages 320–331. Springer.

Lukin, S. M., Reed, L. I., and Walker, M. A. (2015). Generating sentence planning variations for story telling. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 188.

Lukin, S. M., Bowden, K., Barackman, C., and Walker, M. A. (2016). A corpus of personal narratives and their story intention graphs. In *Language Resources and Evaluation Conference (LREC)*.

Moreno, R., Mayer, R., and Lester, J. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. In *World Conference on Educational Media and Technology*, volume 2000, pages 776–781.

Neff, M., Kipp, M., Albrecht, I., and Seidel, H.-P. (2008). Gesture modeling and animation based on a probabilistic re-creation of speaker style. *ACM Transactions on Graphics (TOG)*, 27(1):5.

Neff, M., Wang, Y., Abbott, R., and Walker, M. (2010). Evaluating the effect of gesture and language on personality perception in conversational agents. In *Intelligent Virtual Agents*, pages 222–235. Springer.

Neff, M., Toothman, N., Bowmani, R., Fox Tree, J., and Walker, M. (2011). Don't scratch! self-adaptors reflect emotional stability. In *Intelligent Virtual Agents*, pages 398–411. Springer.

Nelson, K. (2003). Narrative and self, myth and memory: Emergence of the cultural self.

Pennebaker, J. W. and Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of clinical psychology*, 55(10):1243–1254.

Rehm, M., Nakano, Y., André, E., Nishida, T., Bee, N., Endrass, B., Wissner, M., Lipi, A. A., and Huang, H.-H. (2009). From observation to simulation: generating culture-specific behavior for interactive systems. *AI & society*, 24(3):267–280.

Rishes, E., Lukin, S. M., Elson, D. K., and Walker, M. A. (2013). Generating different story tellings from semantic representations of narrative. In *Interactive Storytelling*, pages 192–204. Springer.

Ryokai, K., Vaucelle, C., and Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of computer assisted learning*, 19(2):195–208.

Suzuki, S. V. and Yamada, S. (2004). Persuasion through overheard communication by life-like agents. In *Intelligent Agent Technology, 2004.(IAT 2004). Proceedings. IEEE/WIC/ACM International Conference on*, pages 225–231. IEEE.

Tolins, J., Liu, K., Wang, Y., Fox Tree, J. E., Walker, M., and Neff, M. (2013). Gestural adaptation in extravert-introvert pairs and implications for ivas. In *Intelligent Virtual Agents*, page 484.

Tolins, J., Liu, K., Neff, M., Walker, M. A., and Fox Tree, J. E. (2016a). A verbal and gestural corpus of story retellings to an expressive virtual character. In *Language Resources and Evaluation Conference (LREC)*.

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

Wang, Y. and Neff, M. (2013). The influence of prosody on the requirements for gesture-text alignment. In *Intelligent Virtual Agents*, pages 180–188. Springer.