A Multimodal Motion-Captured Corpus of Matched and Mismatched Extravert-Introvert Conversational Pairs

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

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

Abstract

This paper presents a new corpus, the Personality Dyads Corpus, consisting of multimodal data for three conversations between three personality-matched, two-person dyads (a total of 9 separate dialogues). Participants were selected from a larger sample to be 0.8 of a standard deviation above or below the mean on the Big-Five Personality extraversion scale, to produce an Extravert-Extravert dyad, an Introvert-Introvert dyad, and an Extravert-Introvert dyad. Each pair carried out conversations for three different tasks. The conversations were recorded using optical motion capture for the body and data gloves for the hands. Dyads' speech was transcribed and the gestural and postural behavior was annotated with ANVIL. The released corpus includes personality profiles, ANVIL files containing speech transcriptions and the gestural annotations, and BVH files containing body and hand motion in 3D. The corpus should be a useful resource for researchers working on generating human-like and adaptive multimodal behaviors in intelligent virtual agents.

Keywords: spontaneous communication, gesture production, personality, motion capture, data gloves, entrainment, nonverbal behavior generation

1. Background and Motivation

This paper presents a new corpus, the Personality Dyads Corpus, consisting of multimodal data for three conversations each between three two-person dyads (a total of nine dialogues across three sessions), two matched for personality and one mismatched. The goal of the corpus is to support research on how personality and entrainment affect conversational behavior.

It is well known that the behavior of people engaged in dialogue is influenced both by their individual personalities and by their joint participation in the collaborative act of communication. However strongly individuality may be conveyed by movement, gesture, and linguistic style, conversational participants dynamically adapt to their conversational partner. Theories that attempt to explain such adaptation include the collaborative theory of language use (Clark, 1996; Schober, 1998), theories of alignment (Branigan, Pickering, Pearson, & McLean, 2010), communication accommodation theory (Tickle-Degnen & Rosenthal, 1990; Giles, Taylor, & Bourhis, 1973; Giles, Coupland, & Coupland, 1991), and narrative theories of identity in interaction (Thorne, 1987; McAdams & Pals, 2006).

Entrainment arises out of interlocutors grounding communication with each other. As a conversation develops, conversational participants collaborate on conversational goals, the meanings they are communicating with each other, what perspective to take in approach to a situation, how to refer to topics in joint attention, and other things. For example, participants may create conceptual pacts on how to refer to an object, agreeing to talk about the same object as either a shoe, a loafer, or a brown loafer (Brennan & Clark, 1996) or show preference for gestures more like their own (Luo, Ng-Thow-Hing and Neff, 2013).

Language and gesture are thought to be closely tied to one another, linked both temporally and semantically (McNeill, 1992), thus gestures are likely to play a role in the collaborative grounding that forms the basis of communication (Holler & Wilkin, 2011). Gestural entrainment is displayed through the repetition of gestural patterns across conversational partners; it has also been labeled imitation, gestural mimicry, and alignment. Studies that have looked beyond just verbal communication have found striking commonalities in adaptation and alignment across a number of modalities co-produced with speech. For example, people imitatively match facial and gestural displays, with imitative behaviors increasing with increased social affiliation (Louwerse, Dale, Bard, & Jeaniaux, 2012). In a large corpus study, Bergmann and Kopp (2012) found evidence for alignment of some, but not all features of co-speech gestures, demonstrating that gestural entrainment is not necessarily an all-or-none replication. Instead, individuals engaged in a conversation may be adapting to each other's nonverbal expressive style, rather than repeating gestures on a one-to-one basis.

Another aspect that must be taken into account and coordinated during spontaneous interaction is the personality of the interactants. Several theories provide a basis for personality adaptation. Based on the notion of *press* (Murray, 1938), these approaches consider the dynamic expression and interaction of personality dispositions over the course of a conversation (see e.g. Thorne, 1987). However, none of these theories specify how the time course of local linguistic or gestural adaptation may be dependent upon the personality traits of both conversational partners. That is, the relationship between personality press and conversational dynamics and adaptation may be mutually influential.

The domain of nonverbal behaviors in which information indicative of personality traits may be couched is large and diverse, including modalities such as facial expressions, posture, and gait, as well as movement of the head, hands, and arms, all of which can be further divided into numerous features. While judgment is most accurate when all modalities are viewed together, including both verbal and nonverbal expressions, observing bodily communication alone does allow for some accurate judgments of the observed individual to be made (Ekman & Friesen 1974). These communication channels have been studied from a number of perspectives including emotion (Walbott, 1998; Waxer, 1977), relationship status (Mehrabain 1969), as well as a variety of different measures of personality (Riggio & Friedman, 1986; Gallaher, 1992; Friedman, DiMatteo, & Taranta, 1980).

Many of these studies looked at how individual features, such as body posture, may be indicative of particular personality traits. Others, however, have taken a more holistic view and have attempted to make use of a large number of different measures of nonverbal behavior in order to reveal underlying patterns of behavior. How we categorize nonverbal communication channels into composite features determines along which dimensions variation may be used to indicate differences in personality traits. Considering gesture as a paradigmatic example, a number of taxonomies have arisen. Gestures can be divided along their temporal structure, through which a single gesture can be described generally as consisting of a preparation phase, a hold phase that occurs before and after the main phase, the main stroke phase of the gesture, and a retraction (McNeill 1992; Kita et al. 1998; McNeill 2005). They may also be divided in terms of type such as iconic, metaphoric, beat, or deictic (McNeill, 1992; Goldin-Meadow, 1999). Lastly, the physical features of the gesture may be categorized including length and speed of the trajectory of the gesture, distance from the body forward and outward, and height (Kipp, Neff, & Albrecht, 2006). These various categorization strategies have provided the basis for the exploration of how personality traits are expressed through individual and stylistic variation in gesture (Riggio & Friedman, 1986; Gallaher, 1992; Neff, Toothman, Bowmani, Fox Tree, & Walker, 2011), and may be of further importance when considering the relationship between personality and adaptation in communicative behaviors across conversational partners.

As the exploration of nonverbal expressive behaviors has developed, the extraversion/introversion scale of the Big 5 model of personality (see e.g. Goldberg, 1996) has become the most well studied, due to its high visibility and agreement across judgments (John & Robbins). Riggio and Friedman (1986) were among the earliest researchers to focus on personality and nonverbal cues. They found a positive correlation between extraversion, measured through the Personality Research Form (Jackson, 1974), and gestural fluency for both males and females, a dimension that includes measures of object-focused and parallel gestures, posture shifts, and smiles, as well as body emphasis for males. Making use of a different set of features of nonverbal behavior, Lippa (1998) correlated judgments of extraversion with judgments of recorded spontaneous behavior for measures that ranged from microscopic (i.e. hands away from body) to macroscopic (i.e. broad gestures) levels. In harmony with previous findings, a positive correlation between extraversion and movements of the hands and elbows away from the body was found.

What role would the personality traits of the conversational partners play in the development of gestural entrainment and alignment? While many personality theories have focused on traits that are stable and replicable, traits that individuals possess, it is also useful to think about personality as a process based on an individual's activities, traits that individuals enact (Cantor, 1990; Harlow & Cantor 1994). By acting in a particular fashion, speakers provide particular opportunities for whomever they happen to be engaged with, actively changing the social context of the interaction (Eaton & Funder, 2003). Interacting extraverts and introverts who were matched on personality had more distinctive conversational styles than those who were not matched, for example two extraverts engaged in conversations that were upbeat and expansive whereas two introverts were more likely to be serious and focused (Thorne, 1987). It's possible that this distinctiveness was seen not only in the types of speech acts and conversational structures measured, but also in the nonverbal expressive behaviors such as gesturing. For example, extraverts were observed to focus more on developing common ground with their conversational partner, so it may be that extraverts are also more likely to actively pursue gestural entrainment. Conversely, extravert-extravert pairs may make use of increasingly extraverted gesturing styles over the course of a conversation, increasing along the dimensions of expansiveness, fluency, and animacy.

Previous work and corpora available for examining entrainment have primarily been collected using controlled tasks, such as referential communication tasks, in which pairs repeatedly label a set of objects together, and direction-giving tasks, in which pairs negotiate routes, often with a director-follower structure and with repetitions of directions across conversational partners. Repetition of directions is particularly useful when considering the re-use of gestures across conversational partners, but narrows the focus to emphasize gestures focused on spatial movement, and similarly focuses on repetition of specific forms over stylistic dimensions. Our goal was to collect conversational behavior in a less controlled setting, focusing on open-ended and abstract dialogue, in order to examine the effect on behavior of both personality and dynamic adaptation to a conversational partner. Rather than looking for one-to-one matches in gesture production, we provide annotated data useful in the exploration of stylistic adaptation, visible regardless of whether or not the form of the gestures are the same.

The conversational partner was either of a similar or a different self-reported level of extraversion, the trait most clearly associated with specific nonverbal expressive behavior. The corpus can help achieve the goal of developing an understanding of how entrainment may vary across dialogue pairings, as well as how global features of gesture style may be adapted over the course of a conversation. By providing detailed information regarding the form and temporal structure of the gestures, this corpus will also be of use to those who wish to incorporate embodied dimensions in systems designed to interact with humans, such as interactive virtual agents. (cf., Liu, Tolins, Fox Tree, Neff, & Walker, 2013, 2016; Tolins, Liu, Neff, Walker, & Fox Tree, 2016). The corpus is available at nlds.soe.ucsc.edu/corpora.

2. Corpus Description

The corpus consists of multimodal data for multiple conversations from three two-person dyads. Participants were recorded using optical motion capture for the body and data gloves for the hands in order to permit a more detailed future analysis examining gesture and hand shape, posture changes, and precision timing. This decision represents a tradeoff. Motion capture imposes an additional cost on the participants as they must wear special suits on top of their clothing, as well as gloves, and go through a calibration process that requires a little over one hour. This means that the total capture session lasted 2-3 hours for a single dyad. The gloves and suits do add a slight encumbrance that may damp down smaller movements. There is also a risk that participants are more self-conscious, but it was not clear in practice that this exceeded the impact of their merely being aware that they were being filmed. There is also a significant cost in terms of cleaning both the body and glove data to fix errors that can arise due to noise, inaccurate calibration, and occluded or mislabeled motion capture markers (markers are captured as 3D locations and must be labeled in a semi-automatic process to specify where they are attached to the body).

2.1 Participants

Fifty-five participants (38 female, 17 male, age range from 18 to 50) were recruited through advertisements in the local newspaper and were asked to complete an online pretest personality survey, the BFI (John, Donahue, & Kentle, 1991), from which a personality profile was calculated measuring the Big Five personality traits (John & Srivastava, 1999). Participants were selected based on their profile and asked to come to the lab for a session. Six participants (three male three female) were chosen who scored at least .8 of a standard deviation above or below the sample mean on the personality profile's extraversion scale. This was done in order to find individuals who rank either highly extraverted or highly introverted, given a particular sample population. We recorded personality-matched and -mismatched pairs: an Extrovert-Extrovert dyad, an Introvert-Introvert dyad, and an Extrovert-Introvert dyad.

2.2 Participant Tasks

Participants completed three conversational tasks. These tasks were designed to provide an opportunity to engage in spontaneous, open-ended conversation. Previous research exploring gesture adaptation has made use of either direction giving (Bergmann & Kopp, 2012) or story retelling (Mol et al., 2012). In contrast, the three interactions for the current study were more abstract in nature, providing a novel conversational format for the analysis of nonverbal expressive alignment. Activity 1 was a get-to-know-you task in which the participants were asked to tell each other three things about themselves, two of which were true and one of which was a lie. The goal of the conversation was to figure out which was the lie. In Activity 2, participants were asked to ask each other questions such as "Who is your favorite superhero and why?," "If you won a million dollars in a lottery, what would you do with it?," or "If you could travel anywhere in the world instantly right now, where would you go?" In the Activity 3, the participants were asked to discuss what kinds of movies, TV shows, or websites they both had a shared interest in. They were then to pick one and discuss how they would summarize and promote that thing to people who have never heard of it before, with the goal of coming up with a description they could provide to someone who was unfamiliar with the topic to convince them of its quality. Between activities, participants were individually questioned by an experimenter for information on how well the conversation had gone and what they remembered from the task. At the end of all three tasks, the participants were asked about their sense of liking and rapport with their conversational partner (Gratch, Wang, Gerten, Fast, & Duffy 2007; Puccinelli & Tickle-Degnen, 2004).

2.2 Data

Participant sessions were recorded using two video cameras, audio, motion capture and hand capture. Motion capture was performed with a Vicon optical motion capture system consisting of 12 4-megapixel cameras, hung nine feet above the ground on rails around the perimeter of the motion capture studio. Participants wore specially designed motion capture suits that fit over street clothing in order to increase their comfort. Reflective markers were attached to these and their 3D locations were recorded by the system and later used to reconstruct a skeleton following the participant's movement. Each participant also wore a pair of CyberGlove data gloves. These used bend sensors to measure hand shape and data is recorded wirelessly (see Wang & Neff, 2013 for details).

In addition to recording behavior with the motion capture system, we recorded behavior using two video cameras, one focused on each interlocutor with a full-body frame. There was a shotgun microphone attached to one of the cameras to get the best possible audio for transcription of the talk produced during the interactions. Figure 1 shows two shots of the same scene in an interaction. Note, for privacy reasons, we will not be releasing the video. We will release the motion capture of the body and hands in BVH format. This file format contains two sections. A preamble defines the skeleton structure used to represent the subject. A data section then contains the angles for each joint in the skeleton at each frame. Frames are sampled at 120 fps. This data can be played in ANVIL, a free transcription software (Kipp, 2010), along with the transcription and gesture annotation.

Dialogue was transcribed using an abbreviated version of the Jeffersonian Transcription System (Psathas & Anderson, 1990). Gestures were transcribed using a three-tiered system based on those developed by Kita and colleagues (1998) and by Kipp and colleagues (Kipp, Neff, & Albrecht, 2007). This gesture transcriptions was developed and conducted making use of the video recordings from the interactions (not the motion capture data). Gesture coding was thus entirely manual. The highest level, the g-unit, captured the total time from the beginning of the moving of the hands to prepare for a gesture or series of gestures until the end. The lowest level, the g-phase, captured the temporal structure of the gesture, separating out the preparation phase, the stroke (the most meaning carrying phase), holds, and retractions. In between, the g-phrase captures complete individual gestures. All annotations were completed in ANVIL.

These measures allowed us to capture a number of traits that have previously been demonstrated to be indicative of an individual's extraversion/introversion. For example, by comparing the number of gestures with the numbers of words produced by individual speakers, we can provide a measure of gesture rate, such that a higher gesture rate represents more gesturing produced over the course of the same number of individual words. Second, the broadness of the gestures was calculated, taken as the average radial arm distance of all gestures for the portion of talk. Outwardness of the gestures was measured as the average distance from the body across the three dimensions coded. Finally, the average angle of the elbow to body was calculated. These four measures – rate, expansiveness, outwardness, and elbows-out-ness – have all been previously shown to be features of gestural style that are positively correlated with extraversion (Gallaher, 1992), and have been successfully implemented in an interactive agent (Neff, Wang, Abbott and Walker, 2010). Measurements were taken for the first and third activity of the session in order to capture change over the course of the interaction.

Figure 1A: Extravert-Extravert pair demonstrating a broad gesture.

Figure 1B: Extravert-Extravert pair demonstrating a broad gesture.

3. Corpus Characteristics

The corpus consists of talk produced by three dyads, each engaging in three separate conversational interactions, totaling 67 minutes of footage total (average interaction length = 7 minutes). The corpus contains 1019 conversational turns, consisting of 9923 total words and 882 total gestures (see Figure 2 for a sample interaction). Gestures were hand coded following the three-tiered annotation scheme described above. See Figure 3 for an example of Anvil structure and Figure 4 and Figure 5 for an illustration of what has been captured by motion capture. These files have detailed annotations and can be played back to examine both gestural and postural behavior and entrainment.

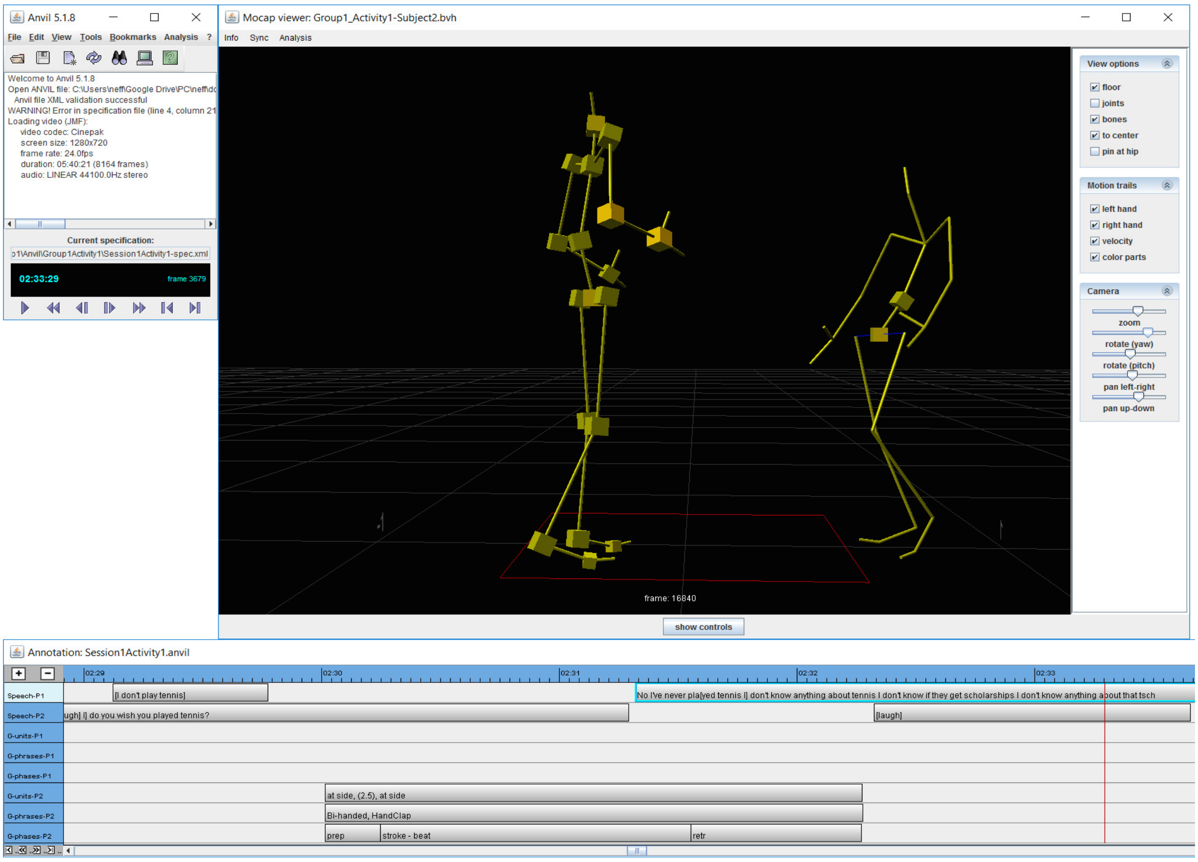

Figure 4: The motion capture files can be displayed in ANVIL as 3D stick people (top middle). The interface allows for rotation, panning, and zooming. The bottom dialog shows the annotation tracks. These contain, in order, the dialog for P1, dialog for P2, gesture Unit, Phrase, and Phase for P1 and gesture Unit Phrase and Phase for P2. The playback control is located in the top left.

P2: So what is true and what is the lie:: (.) so you said you (.) lemme see if I got what you said
P1: Mhm
P2: U::h you played tennis but I don't remember (.6) cause you were talking fast at the start
P1: Ohh yeah I play tennis with UC Davis tennis team
P2: UC Davis tennis team.
P1: Mhm
P2: You were here on a scholarship?
P1: No chance (.) doesn't have any scholarships for tennis
P1: Yeah. It's not like the- like there's football and that's like a bi::g sport
P2 Right Right
P1: Baseball you get a little I- I used to date a baseball player and he gets a lot of money not a lot of money but more money than we get it's unfair
P2: Yeah

Figure 2: Dialogue transcription taken from the first activity of the Extravert-Extravert paired dyad. (.) represents pauses, colons represent elongated vowels, and dashes represent words that are cut short.

Previous analyses using similar data demonstrates avenues of research made possible by such a data set. In Tolins et al. (2013), an Extravert-Extravert and Extravert-Introvert dyad were compared, providing a descriptive analysis of the nonverbal expressive behaviors of the two dyads as they interacted across the three tasks. The measures used to describe the correlates of extraversion in nonverbal expressive style (Gallaher, 1992) indicate that all participants' styles changed over time. Across measures such as gesture rate, gesture broadness and expansiveness, and elbow distance from the body, the data suggested two things: (1) that people adapt to their interlocutor and (2) that personality may modulate this adaptation.

In the extravert dyad, along two gesture traits each of the pair changed their style of gesture in the same direction (Tolins et al., 2013). Along the scales of elbow expansion and broadness, both participants became more stylistically extraverted over time. For two other traits, the members of the extravert dyad changed to be like each other. Along the scale of rate, one extravert reduced rate to match the other extravert. Along the scale of

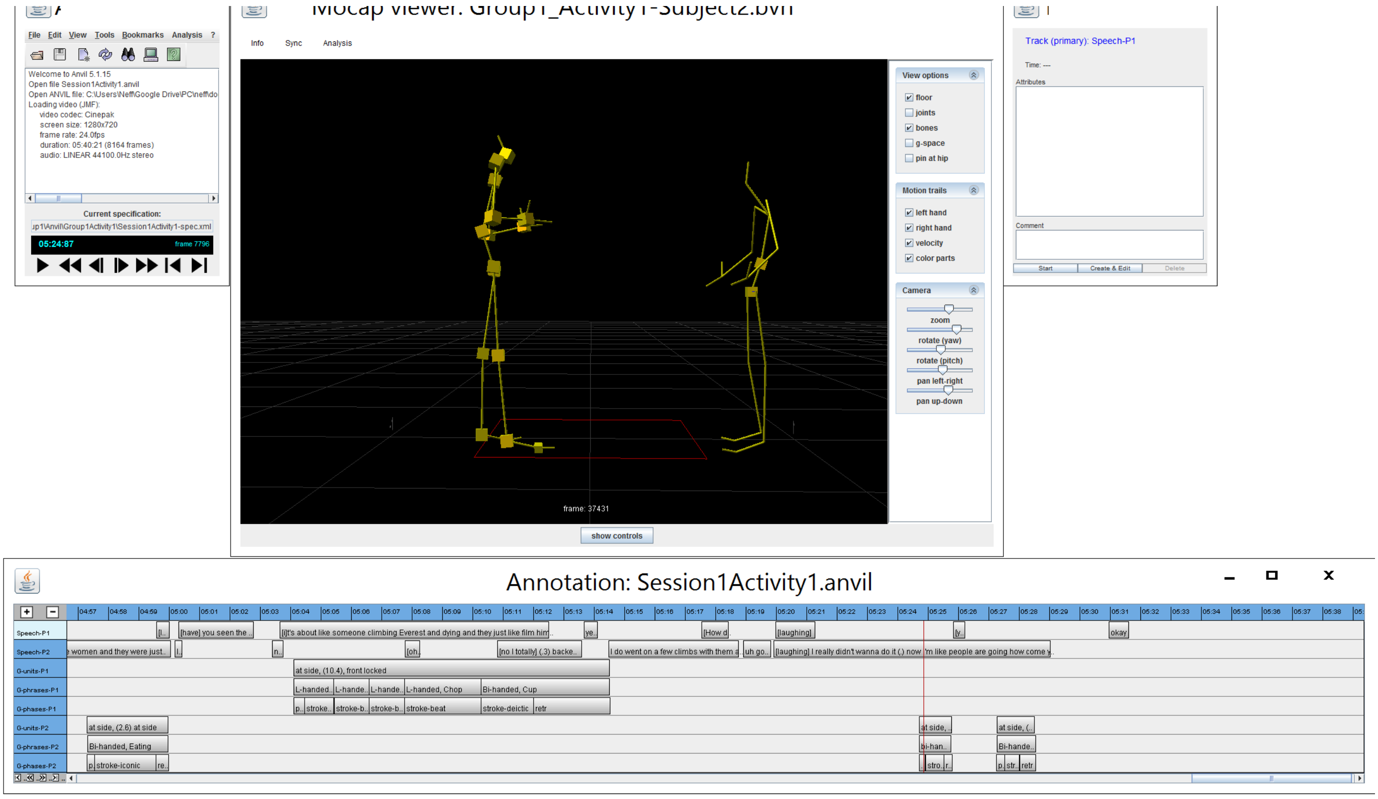

Figure 5: A second example illustrating the time course of the motion capture files. These files are displayed in ANVIL as 3D stick people (top middle). The interface allows for rotation, panning, and zooming. The bottom dialog shows the annotation tracks. These contain, in order, the dialog for P1, dialog for P2, gesture Unit, Phrase, and Phase for P1 and gesture Unit Phrase and Phase for P2. The playback control is located in the top left.

outwardness, each member of the extravert dyad moved towards the other. In the mismatched pair it was the extravert who demonstrated the greatest shift in style. This movement was towards a more extraverted style for gesture rate, as the extravert adapted to the higher-rate introvert. It was towards a more introverted style with elbow expansiveness, as the extravert adapted to the less-expansive introvert. With broadness, the introvert became broader and the extravert became narrower. For outwardness, both members of the extravert-introvert dyad reduced outwardness.

While all participants in Tolins et al. (2013) changed their gestures over time, those individuals who were more extraverted were more apt to shift the features of their gestures to match their partner. Importantly, this adaptation was measured in stylistic, or non-communicative, features of gesture, and would be considered a type of mimicry or alignment. Gestural alignment has typically been found in tasks involving giving and receiving directions, or taking turns describing objects, with a focus on the form of the gesture. We investigated gesture produced in an open, spontaneous dialogue on an abstract subject, and measured partner-specific adaptation in terms of stylistic dimensions of expressive behavior.

We note that personality is only one variable that may affect the production of nonverbal behavior, including how nonverbal behavior changes over time. Other factors that may affect behavior include both the rapport established across conversational partners as well as individual emotion and mood (André et al. 2000;; Cassell et al. 2007; Zhao et al, 2014; Gratch et al. 2007). These factors may be involved both in shaping the production of gestures as well as influencing the degree to which an individual adapts to their conversational partner.

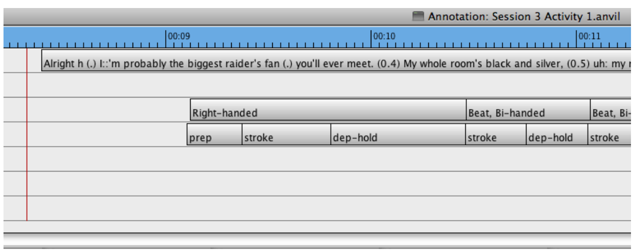

Figure 3: Screen capture of speech and gesture transcription in ANVIL, demonstrating the speech and gesture phase and phrase transcription structure.

There are likely many reasons why one person adapts to another and another person does not. As with entrainment, other factors might include power dynamics between the interactants, prior shared history, and what interactants think of themselves and others.

4. Conclusion

Measures of nonverbal expressive style indicate that gesture correlates are not stable over time, but rather fluctuate in a way that appears to be motivated by the conversational partner (Tolins et al., 2013). While more data is needed before confident generalization to a larger population can be made, the existing data set is useful for researchers seeking detailed information about how some people with particular personality profiles on the extreme of the extraversion-introversion scale behave across a set of dialogues over the course of one recording session.

This corpus could be useful for researchers investigating how personality affects dyadic interaction on verbal and bodily levels. Close investigations could include looking at how gestures that are co-produced with particular words, syntactic structures, or emotional content vary based on personality or time. See Hu et al. (2016) for a demonstration of a possible implementation based on such a paradigm. For example, a gesture used to identify a referent could be quickly taken up by an addressee of a matched personality, but slowly taken up by an addressee of a mismatched personality. As another example, the ratio of beat, iconic, metaphoric, and deictic gestures may vary for different personality types, or as people interact with each other over time. People may also display different postural changes based on their personality or time. As was observed with gestural adaptation in Tolins et al. (2013), postural adaptation may also look different between matched and mismatched dyads, or from earlier versus later in the recording session.

Another possible use of this corpus is in developing character models for virtual agents, particularly in dyadic interaction. The corpus can also be useful for creating virtual agents that can interact with people with differing personality types (Kopp et. al 2006; Gratch & Marsella 2001; Häring et al. 2011; Van Mulken et al, 1998; Rehm et al. 2009; André al. 2000). These agents would gestures much like another human would, evoking a genuine sense of interaction through relatively straightforward gestural cues, thereby increasing rapport and social affiliation (Cassell et al. 2007; Zhao et al, 2014; Gratch et al. 2007). An agent may appear robotic and unconvincing if it does not take the human interactant's gestures into account and adapt to them. People both notice adaptation in agents and tend to prefer agents that adapt to each other (Hu et al., 2015). Creating realistic models of how agents with different personalities respond to each other's behavior will require corpora like ours that provide precise details about how people with different personalities respond to each other, both linguistically and bodily.

The recording of high quality motion capture data opens up two further avenues of use for the corpus. First, it allows very precise, fine-grained analysis of character pose and relative positioning throughout the conversations. The 3D locations of any joints can be easily calculated, along with the joint angles. Other quantities, like the center of mass and balance point can be reasonably approximated. This allows a level of detail in the analysis that is not possible with video data. Second, the motion capture data can be directly used in the range of animation applications that take motion capture as their input data. For example, motion graphs (Kovar et al., 2002) can be built from the segments, allowing new conversational sequences to be generated. As another example, splicing could be used to apply limb motion to different motion sequences (Heck et al. 2006; van Basten & Egges, 2012). Data could also be used in perceptual studies (cf. Jörg et al., 2010; Wang & Neff, 2013; Hoyet et al., 2013).

5. Acknowledgements

This work was supported by NSF Grants IIS-1115742 and IIS-1115872. We would like to thank the many research assistants who were involved in the coding of the corpus.

6. References

André, Elisabeth, Martin Klesen, Patrick Gebhard, Steve Allen, and Thomas Rist. "Integrating models of personality and emotions into lifelike characters." In *Affective interactions*, pp. 150-165. 2000.

Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108–121.

van Basten, B., & Egges, A. (2012). Motion transplantation techniques: A survey. *Computer Graphics and Applications, IEEE*, *32*(3), 16-23.

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between humans and computers. *Journal of Pragmatics, 42,* 2355-2368.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22,* 1482-1493.

Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. (1997) 415–438.

Cassell, J., Gill, A. J., & Tepper, P. A. (2007). Coordination in conversation and Rapport. In *Proc. Workshop on Embodied Language Processing* (pp. 41-50). ACL.

Clark, H. H. (1996). *Using Language*. New York: Cambridge University Press.

Gallaher, P. E. (1992). Individual differences in nonverbal behavior: Dimensions of style. Journal of personality and social psychology, 63(1), 133-145.

Giles, H., Coupland, J., & Coupland, N. (1991). Acommodation theory: Communication, context, and consequence. In Giles, H., Coupland, J., & Coupland, N. (Eds.), *Contexts of Accommodation.* New York, NY: Cambridge University Press.

Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: Some Canadian data. *Lang Soc., 2,* 177-192

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating rapport with virtual agents. In *Intelligent Virtual Agents* (pp. 229-238). Springer Berlin Heidelberg.

Gratch, J., & Marsella, S. (2001). Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In *Proc. of the 5th international conference on Autonomous agents* (pp. 278-285). ACM.

Häring, M., N. Bee, and E. André (2011)."Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots." In *Ro-Man, 2011 IEEE* pp. 204-209.

Hoyet, L., Ryall, K., Zibrek, K., Park, H., Lee, J., Hodgins, J., & O'Sullivan, C. (2013). Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies. *ACM Transactions on Graphics (TOG)*, *32*(6), 204.

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

Hu, Z., Dick, M., Chang, C.-N., Bowden, K., Neff, M., Fox Tree, J. E., Walker, M. (2016). A corpus of gesture annotated dialogues for monologue to dialogue generation from personal narratives. *Proc. of the Language Resources and Evaluation Conference (LREC)*.

Jackson, D. N. (1974). *Personality Research Form Manual.* Port Huron: Research Psychologists Press. Rev. ed.

Jörg, S., Hodgins, J., & O'Sullivan, C. (2010, July). The perception of finger motions. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (pp. 129-133). ACM.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The "Big Five" Inventory – V. 4a and 54.*: UC Berkeley Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In: *Handbook of Personality: Theory and Research* (pp. 102-138).

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. (2010). Multimedia annotation, querying and analysis in ANVIL. Multimedia information extraction, 19.

Kita, S., Van Gijn, I., & Van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In: *Gesture and Sign Language in Human-Computer Interaction (pp. 23-35).*

Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., & Vilhjálmsson, H. (2006). Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent virtual agents* (pp. 205-217).

Kovar, L., Gleicher, M., & Pighin, F. (2002, July). Motion graphs. In *ACM transactions on graphics (TOG)* (Vol. 21, No. 3, pp. 473-482). ACM.

Liu, K., Tolins, J., Fox Tree, J. E., Walker, M., & Neff, M. (2013). Judging IVA personality using an open-ended question. In *Springer's Lecture Notes in Computer Science/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.

Luo, P., Ng-Thow-Hing, V., & Neff, M. (2013, January). An examination of whether people prefer agents whose gestures mimic their own. In Intelligent Virtual Agents (pp. 229-238). Springer Berlin Heidelberg.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist, 61,* 204-217.

Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn.

Neff M., Wang Y., Abbott $. & Walker, M. (2010). Evaluating the effect of gesture and language on personality perception in conversational agents. In Intelligent Virtual Agents (pp. 222-235).

Puccinelli, N. M., & Tickle-Degnen, L. (2004). Knowing too much about others: Moderators of the relationship between eavesdropping and rapport in social interaction. *Journal of Nonverbal Behavior, 28,* 223-243.

Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica, 78,* 75-99.

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

Schober, M. (1998). How addressees affect spatial perspective choice in dialogue. In: P. L. Olivier & K. P. Gapp (Eds.), *Representation and processing of spatial expressions* (pp. 231-245).

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1,* 285-293.

Thorne, A. (1987). The press of personality: A study of conversations between introverts and extraverts. *Journal of Personality and Social Psychology, 53,* 718-726.

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 *Springer's Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence* LNAI 8108 (pp. 481–482).

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

Van Mulken, S., E. André, and J. Müller (1998). The persona effect: how substantial is it."*People and computers XIII: Proc. of HCI* 98: 53-66.

Wang, Y., & Neff., M. (2013) Data-driven glove calibration for hand motion capture. *In: Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation.*

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

Zhao, R., Papangelis, A., & Cassell, J. (2014). Towards a dyadic computational model of rapport management for human-virtual agent interaction. In *Intelligent Virtual Agents* (pp. 514-527).