View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Union College: Union | Digital Works

Union College
Union | Digital Works

Honors Theses

Student Work

6-2016

Effect of Deictic Gestures on Direction-Giving in Virtual Humans

Anthony Pham
Union College - Schenectady, NY

Follow this and additional works at: https://digitalworks.union.edu/theses

Part of the Graphics and Human Computer Interfaces Commons, and the Interpersonal and Small Group Communication Commons

Recommended Citation
Pham, Anthony, "Effect of Deictic Gestures on Direction-Giving in Virtual Humans" (2016). *Honors Theses*. 197.
https://digitalworks.union.edu/theses/197

This Open Access is brought to you for free and open access by the Student Work at Union | Digital Works. It has been accepted for inclusion in Honors Theses by an authorized administrator of Union | Digital Works. For more information, please contact digitalworks@union.edu.

Effect of Deictic Gestures on Direction-Giving in Virtual Humans

Anthony Pham

May 11, 2016

Abstract

Virtual agents are animated characters that use speech and gesture to interact with human users. They can serve as an intuitive interface for a variety of purposes. I am investigating the use of deictic gestures by a direction-giving agent. Deictic gestures are pointing gestures that humans often use in direction-giving to help clarify the route and destination.

In my experiment, I developed a virtual agent to give directions to people to six different locations with the following spatial relationships to the starting point: left, right, left behind, right behind, left up, and right up. Three versions of the virtual agent were programmed to give directions to these locations. The versions differed in their use of gesture: no gestures; only body orientation and head gestures; and body orientation, head gestures, and arm movements. Participants received directions from my agent to six different locations using different gesture modes. Reception of the virtual agent was recorded through a survey and video recordings of the participants' reactions to the direction-giving, as well as the path the participants took to follow the directions. The data was analyzed to assess the agent's effectiveness. Upon analysis, the higher levels of deictic gesture appeared to be more effective, but data collected did not have statistical significance.

Contents

List of Figures

List of Tables

1 Introduction

A virtual agent is a digital entity programmed to help people with different tasks. For example, refer to Figure 1 for an example of a virtual agent that helps users answer shipping-related questions. Virtual agents often interact with people by giving them helpful information. It is important to ensure that the information a virtual agent provides is successful in helping out people. Virtual agents are important as an interface for the exchange of information because they provide a way to process information though an interactive and social manner. Virtual agents can vary in how they interact with people; for instance, a virtual agent used as a banking assistant does not need to behave like a direction-giving agent. I plan to use a virtual agent that interacts through gesture and dialogue from a display to give directions to various locations within the college's campus.

A large part of what makes up human-human interaction is non-verbal communication. Gestures, arm movements, eye gazes, and body orientation are all observable forms of non-verbal communication. These non-verbal forms of communication enhance how people interact, often by adding redundancy to the messages being conveyed. For instance, Cassell [1] states that facial expressions change continuously, ″are synchronized to what is going on in concurrent conversation,″ and ″can help disambiguate what is being said.″ Additionally, gestures help listeners ″form a mental representation of the communicative intent of the speaker.″ These forms of non-verbal communication may be applied to human-virtual character interactions to make a virtual character into an informative social being.

Gestures that are frequently used in human interaction include pointing gestures that reference places, objects, etc. in the gesturer′s environment, also known as deictic gestures. Deictic gestures are frequently used between people when giving directions. These gestures help facilitate direction-giving by providing clarity to what is being said by directly referencing it during communication. In order for a gesturing virtual agent to convey information accurately and successfully, it must carry out the effective kinds of deictic gestures when interacting with people.

The motivation behind this research is to study the efficacy of deictic gestures in direction-giving between people and virtual agents. I have developed a virtual character to provide directions to locations around the Union College campus. Directions were be given from a defined area and were be paired along-

Figure 1: Example of a Virtual Agent that helps with shipping-related problems

side deictic gestures. I believe this is important towards future use as interactive displays become more common in public areas. A virtual character that acts naturally with the user will provide a better experience for the user by making them feel more comfortable, which leads to a better understanding when receiving information. I hypothesize that virtual agents that use more gestures in direction-giving are more successful at guiding people to their destinations.

2 Past Research on Virtual Agents and Gestures

Louwerse and Bangerter [4] looked into the relation between gestures and linguistic expressions. Concrete deictic gestures - directly referencing an object using a gesture - was the focus of their study. They argued that gesture and language are intrinsically linked. Concrete deictic gestures map easily to deictic expressions like ”here” and ”there”. Results support that a feature description accompanied by either a deictic gesture or a deictic expression increases accuracy of target identification. However, a use of both deictic gesture and deictic expression presents no additional gains in accuracy. Gestures support joint visual attention with the user and this alignment of gesture to speech has an important impact on the user. With my study I plan to investigate the efficacy of gestures within a virtual agent system. Specifically, I want to determine whether gestures aid in communication between a virtual agent and people as they do between humans.

Evers, *et al* [2] researched body orientation during direction-giving in a 3D virtual environment by a virtual guide. Their experiment has shown that the body orientation of a virtual guide between 180 degrees, directly facing opposite to the participant, and 120 degrees, where the virtual guide adapts their orientation to the participant, had no particular effect on the participant's ability to navigate the virtual environment. While they hypothesized that the 120 degree orientation would be more natural, the participants perceived the 180 degree orientation to be more natural, because they expected the agent to face them while giving directions. Although the researchers did not find much impact in body orientation, they attribute the lack of difference to complicated directions that may have forced the participants to focus more on the verbal aspect of the direction-giving. In my study, I will be expanding on their work by looking at natural transitions in body orientation during direction-giving, such as those associated with initially facing a person and later turning to indicate the direction in which the person should go.

Dai, *et al* [3] did a study on direction-giving between a social robot, an embodied agent, and a GPS map + speech system. They discovered that users will respond more positively to embodied agents than a standard GPS map + speech system. Both the social robot and embodied agent were perceived as more present than the GPS map + speech system, but the embodied agent was viewed as the most natural. A carefully designed version of their embodied agent that used gestures in giving directions from the user's perspective was shown to have better results in the user's remembering and being able to repeat directions, compared to the GPS alone. Additionally, they were able to construct a cognitive map of the directions given by the gesturing embodied agent. This study shows the merits of using a capable virtual agent for purposes such as direction-giving.

These researchers have worked on studying and improving human-virtual character interaction as well how gestures are important in clarifying one's dialogue. The groundwork necessary for a virtual agent to be successful is established in these past works. I aimed to continue this work by researching what specific kinds of deictic gestures should be used for better direction-giving. Thus, I will need to create my agent in a way that will make it appear as a social being.

Gesture Level	Gesture Description
G1	No Gesture
G2	Body Orientation and Posture
G3	Body Orientation, Posture, and Arms

Table 1: Gesture level and description of how the agent interacts at that level.

Spatial Relationship	Location
Left	SE 006B
Right	Wold 031
Left Behind	SE S013
Right Behind	Olin 115/111
Left Up	SE S109
Right Up	Wold 128

Table 2: Spatial relationships and their locations in relation to the starting point from the CROCHET Lab.

3 Designing a Direction-Giving Experiment

To investigate deictic gestures in direction-giving, a participant will interact with a virtual agent I developed, Rachel, in an experiment. In the experiment the participant is given directions to locations from the CROCHET lab delivered by Rachel acting from a display. To investigate the effects of deictic gestures on direction-giving, I looked into three different versions of the virtual agent, each exhibiting different levels of deictic gestures. One version of the agent used no gestures in their direction-giving (G1), another used only body-orientation and posture while giving directions (G2), and the final version used body-orientation, posture, and arm movements (G3). Table 1 shows the gesture levels described. Refer to Figure 4 for a picture of Rachel giving a direction to go left using body-rotation and posture (G2) in the left image and Rachel giving a direction to go left using body-rotation, posture and arms (G3) in the right image. Each version of the agent gave directions to locations in the campus from the CROCHET lab. Every location is associated to one of the six following different spatial relationships to the CROCHET lab: left, right, left behind, right behind, left up and right up. Refer to Table 2 for an overview of the spatial relationships and their specific locations and Figure 2 for a map representation. By having a combination of these gesture and location types, I aim to see the difference in how the participant will gauge the virtual agent and how successful the

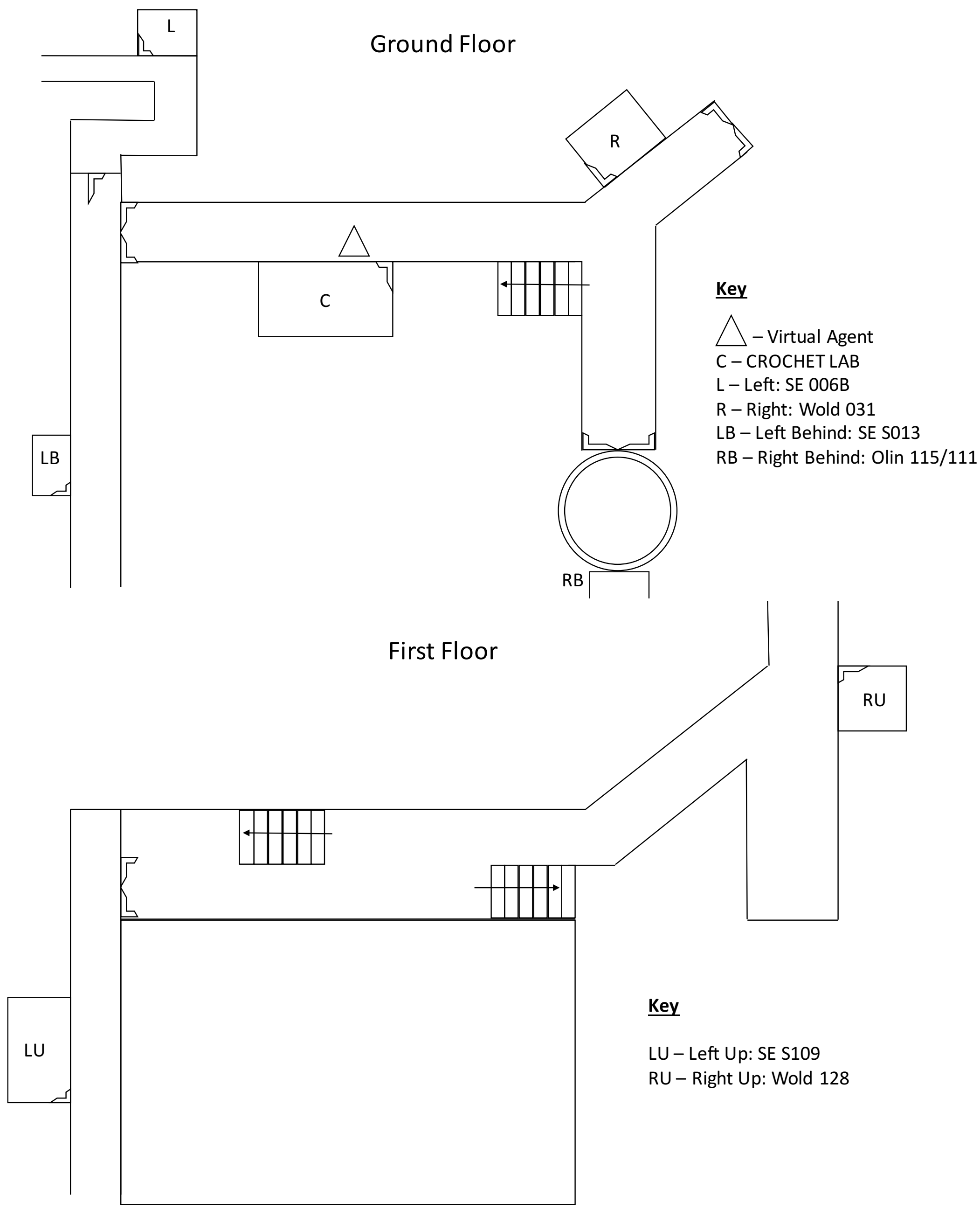

Figure 2: Spatial map of locations in relation to the virtual agent.

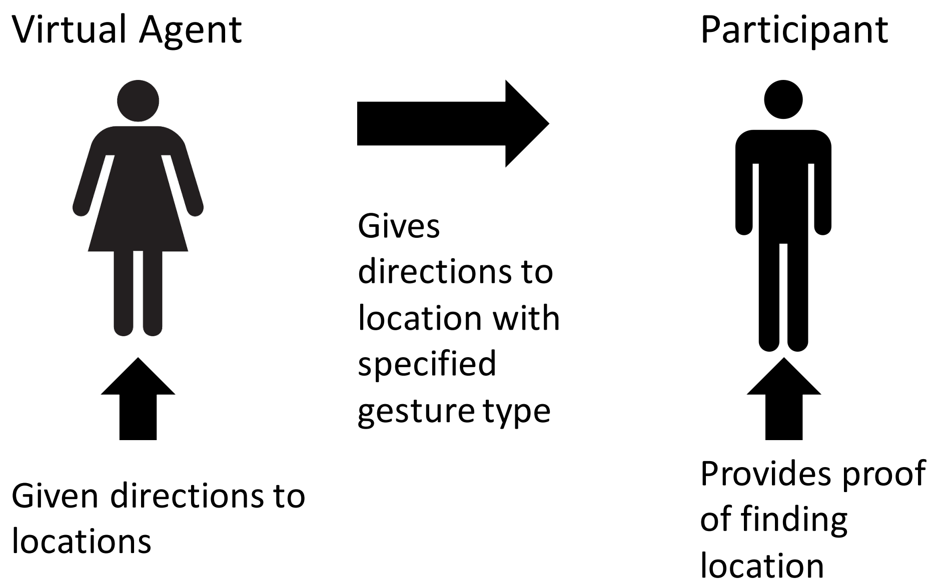

Figure 3: Experiment set-up with Virtual Agent giving directions to a participant.

participant is in finding the given location.

To carry out the experiment, a station was set up outside the CROCHET lab such that the participant could receive directions from the displayed virtual agent. A camera was set up to record the participants' reaction to the direction-giving. The participants were given directions one-at-a-time to six locations based on the spatial relationships. Each set of directions was given by versions of Rachel using different gesture levels. To avoid bias, locations and gesture levels were paired via Graeco-Latin squares such that no set of directions would have the same pair of location and gesture type, and no two participants would receive the directions in the same order. Each participant, then, attempted to find two locations for each gesture type. Refer to Figure 5 for an example of how directions differ between participants. It should be noted that Rachel was set up facing away from the CROCHET Lab; thus, her initial instructions to the participants are from a point-of-view perspective where she is telling the participants where to go based on her spatial relationship to that location. The deictic gestures Rachel uses initially are directly referencing the direction in which the participant must go. After the initial instruction, Rachel gave directions from a route perspective,

Figure 4: G2 version of Rachel gesturing left and G3 version of Rachel gesturing left.

that is, directing as if she were the participant following the directions.

At each location there was a letter posted nearby as a marker for participants to know that they found the right location. The participant was instructed to find a location from the directions given by the agent twice before following the directions to find and write down the letter at the location. If the participant was unable to find the location, they were allowed to return and watch the instructions from Rachel one more time, before they had to seek out the letter again. If the participant was unable to find the location after this, they were instructed to write down a room number near the location they thought Rachel was trying to lead them to. Once the participant found some location, they rated the directions they received on a survey. They were asked to give two ratings of the direction they were given based on helpfulness and naturalness between 1 and 10, with 1 being a low rating and 10 being a high rating. At the end of the experiment, the participants filled out a short response on what they thought the virtual agent did that was helpful and not helpful in the direction-giving. Refer to the appendix for the survey. This repeated for all locations per participant. The participant wore a camera to record the path they took when following directions. Refer

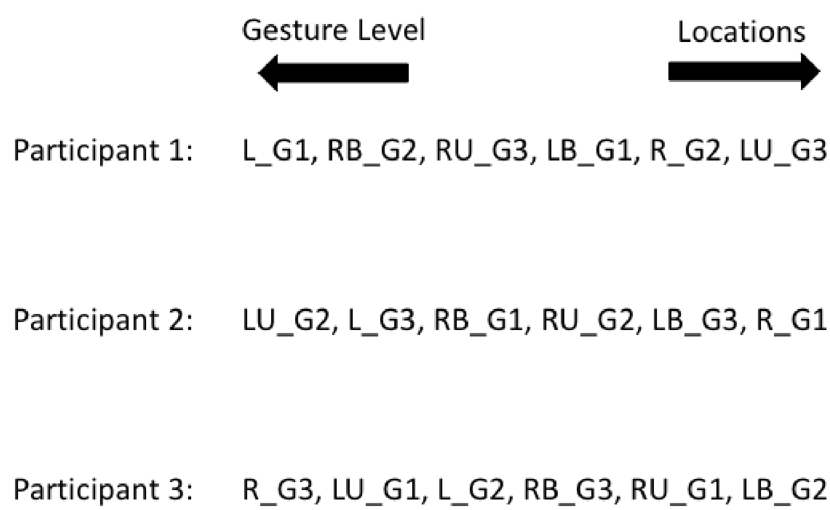

Figure 5: Examples of how directions are given differently between participants using Graeco-Latin squares.

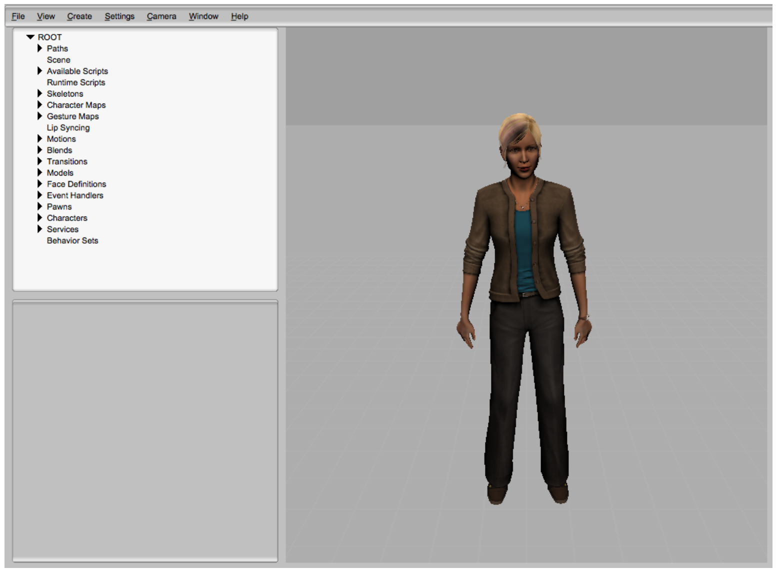

Figure 6: SmartBody GUI with Rachel instantiated

to Figure 3 for a graphical depiction of the experiment.

4 Building a Virtual Agent

I developed my virtual agent, Rachel using SmartBody [5], a character animation platform developed by the USC Institute for Creative Technologies. SmartBody is a Behavior Markup Language (BML) realization engine that transforms BML behavior descriptions into realtime animations. BML commands allow Rachel to execute gaze, gesture, head movements, and use speech to model behavior during direction-giving. BML commands also allowed me to focus on deciding the behaviors of Rachel during direction-giving and coordinating them together, rather than focusing on behavior realization. Refer to Figure 6 for a look at the SmartBody GUI. To build Rachel for my experiment, she will not be functioning autonomously, but will be given pre-programmed directions to the participant. Refer to the appendix for a look at how the directions were coded and scripted for the virtual agent.

Gesture Level	Success	Failed
G1	32	8
G2	40	0
G3	39	1

Table 3: Number of successful direction givings in participants based on gesture level

Averages	Helpfulness	Naturalness
G1	8.3	7.8
G2	8.7	8.1
G3	8.7	8.4

Table 4: Averages of participants' ratings on the directions from the agent on a scale from 1 to 10

5 Results

5.1 Path Finding

The experiment ran for 20 participants. This totals up to 120 direction-giving interactions with my virtual agent, with 40 interactions per gesture level. Most participants were able to find the location on the first try. In 9 of the direction-givings, the participants were unable to find the location based on the directions from the agent. 8 of these were from the version of the agent that used no gestures. Refer to Table 3 for a representation of this data.

5.2 Survey Ratings

Participants were asked to rate the helpfulness and naturalness of the directions they received from the agent after they found the locations during the experiment. In Table 4, the averages are calculated for all the locations across the gesture levels. None of the differences between the averages are statistically significant under a paired t-test at the .05 level, but the p-value when comparing the naturalness of G1 and G3 directions is 0.054.

Averages	Nods	Turns
G1	0.42	0.89
G2	0.49	0.89
G3	0.25	1.02

Table 5: Averages of participants' nods and turns done when actively receiving directions from Rachel

Averages (seconds)	G1	G2	G3
Left	47.5	34.8	49.9
Right	18.0	16.3	18.2
Left Behind	36.3	38.6	37.6
Right Behind	45.4	45.0	65.0
Left Up	51.8	60.4	51.5
Right Up	61.3	65.0	42.6

Table 6: Average time participants took to travel to the locations arranged by location and gesture level

5.3 Survey Responses

At the end of the experiment, participants were asked fill out a short response on what they thought the virtual agent did that was helpful in the direction-giving and what was not helpful. 10 out of 20 participants stated that some form of body motion or gesture was helpful in the direction-giving. Refer to the appendix for all the participants' responses.

5.4 Nods and Turns

Refer to Table 5 to see the average number of nods and turns participants did in reaction to receiving directions from Rachel. Data was unable to be tested for significance as some data was lost in video recordings.

5.5 Time

Refer to Table 6 to see the average time participants took to find the location. Only direction-givings in which the participant was able to find the location without having the directions repeated were considered. Data was unable to be tested for significance as some data was lost in video recordings.

6 Discussion and Conclusion

Although most results gathered from the experiment are statistically insignificant, the trends they follow support my hypothesis that the virtual agent that uses the greatest level of gesture is better overall at direction-giving. In the path-finding results, participants had more trouble following the directions when the virtual agent used no gesture in the direction-giving, with 8 direction-givings being unsuccessfully completed. This is a strong indication of the relative weakness a non-gesturing virtual agent has compared to a gesturing agent in terms of direction-giving. However, the number of successes in G3 compared to G2 differs by only one direction-giving, indicating that the difference in effectiveness between complex gestures vs. simpler gestures is less than the difference between gesture vs. no gesture.

In the survey ratings, helpfulness and naturalness increased as the level of gesture from the agent increased. In this data, the result closest to significance was between the ratings in naturalness in G1 and G3; this indicated that the level of gesture used by the virtual agent is important in the participant's perception of the agent itself. This trend is further reflected in the longer survey responses the participants completed at the end of the experiments. Half of the participants indicated that when the agent used forms of body motion and gesture, they found the directions to be more helpful. The survey results suggest that participants prefer virtual agents that use more complex gestures in G3, because ratings of helpfulness and naturalness were higher than G1 and G2.

The nod and turns and the average time data have reached relatively no conclusion. Along with some of the data having to be discounted, the averages seem to be erratic. For example, when looking at the average number of nods across gesture levels, we see an average of 0.42 nods in G1 that increases to 0.49 nods in G2, but drops to 0.25 in G3. This may be attributed to participants understanding the directions more as the agent gestures, but it is too difficult to tell. It should be noted that the number of nods and turns are relatively low, as most participants only watched the agent give directions with most of the averages being contributed to participants who actively moved when they received directions from the virtual agent. When looking at the averages across all locations and gesture levels for average times, the data jumps all over, making it unreasonable to draw any definitive conclusions. This may be a result of outliers in participants who had particular trouble with direction-giving in general.

Even though the data in its current state is insignificant, it is possible that this positive trend can be reflected in other types of data collection or may become significant if the number of participants was increased. Looking at only 20 participants did not draw enough data to lead to a statistical conclusion, however, having more participants might. Different forms of data collection could also be explored, such as making the surveys more targeted. Many participants rated relatively close on helpfulness and naturalness on the survey; if participants had a targeted question such as, rating one set of directions compared to a different set of directions, it might produce more varying ratings and results. The locations chosen were also relatively close to the CROCHET Lab; perhaps if different locations were explored that required more complex directions and therefore more complex deictic gestures, one might be able to observe that more gesturing is beneficial to direction-giving virtual agents.

References

[1] Justine Cassell et al. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. *Embodied conversational agents*, 1, 2000.

[2] Mark Evers, Mariët Theune, and Joyce Karreman. Which way to turn?: guide orientation in virtual way finding. In *Proceedings of the Workshop on Embodied Language Processing*, pages 25–32. Association for Computational Linguistics, 2007.

[3] Dai Hasegawa, Justine Cassell, and Kenji Araki. The role of embodiment and perspective in direction-giving systems. In *AAAI Fall Symposium: Dialog with Robots*, 2010.

[4] Max M Louwerse and Adrian Bangerter. Focusing attention with deictic gestures and linguistic expressions. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, volume 23, pages 1331–1336, 2005.

[5] Marcus Thiebaux, Stacy Marsella, Andrew N Marshall, and Marcelo Kallmann. Smartbody: Behavior realization for embodied conversational agents. In *Proceedings of the 7th international joint conference on*

Autonomous agents and multiagent systems-Volume 1, pages 151–158. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

7 Appendix

```
bml.execBMLAt(1,'ChrRachel', '<speech type="application/ssml+xml" id="myspeech"> \
<mark name="T0"/>"go" \
<mark name="T1"/> \
<mark name="T2"/>"to" \
<mark name="T3"/> \
<mark name="T4"/>"the" \
<mark name="T5"/> \
<mark name="T6"/>"left" \
<mark name="T7"/> \
</speech> \
<head type="NOD" start="myspeech:T0"/> \
<gaze direction="LEFT" sbm:joint-range="HEAD CHEST" start="myspeech:T6" \
target="ChrRachel"/> \
<gesture name="ChrBrad@Idle01_PointLf01" stroke="myspeech:T6"/>')
```

The code above is a snippet of a BML command that has Rachel say, ”go to the left.” In the bottom half of the command, Rachel is specified to nod at T0, which corresponds with when she says ”go.” When Rachel says ”left,” she will change her gaze to face left using her head and chest and use a gesture to point left.

Gesture Script

Anthony Pham

Left: Biology Faculty Research Lab S006A

G1: No Gesture

Go to my left

Walk through the double doors

Go down the hallway and go through the door on your right

Follow the hallway and it's straight ahead

G2: Body rotation/posture

Go to my left <*Turn body slight left and head tilts left*>

Walk through the double doors

Go down the hallway and go through the door on your right <*Turn body slight right and head tilts right*>

Follow the hallway and it's straight ahead

G3: Body + Arms

Go to my left <*Turn body slight left, head looks left, left arm extends, and left hand is pointing to the left*>

Walk through the double doors

Go down the hallway <*Both arms up at 90 degrees and extend to 180 degree angle*> *and go through the door on your right* <*Head tilts right and right hand points to the right*>

Follow the hallway and it's straight ahead

Right: Music Lab

G1: No Gesture

Go to my right

At the end of the stairs turn slight left and the location will be to your left

G2: Body rotation/posture

Go to my right <*Turn body slight right and head tilts right*>

At the end of the stairs turn slight left <Turn body slight left>

and the location will be to your left <Head tilts left>

G3: Body + Arms

Go to my right <Turn body slight right, head looks right, right arm extends, and right hand is pointing to the right>

At the end of the stairs turn slight left <Turn body slight left and head tilts left> and the location will be to your left <Left hand is pointing to the left>

Left Behind: S&E Basement Elevators

G1: No Gesture

Go to my left

Walk through the double doors

Go down the hallway and turn left

Follow the hallway and turn right then left

The location will be to your right

G2: Body rotation/posture

Go to my left <Turn body slight left and head tilts left>

Walk through the double doors

Go down the hallway and turn left <Turn body slight right and head tilts right>

Follow the hallway and turn right <Head tilts right> then left <Head tilts left>

The location will be to your right <Turn body slight right and head tilts right>

G3: Body + Arms

Go to the my <Turn body slight left, head looks left, left arm extends, and left hand is pointing to the left>

Walk through the double doors

Go down the hallway <Both arms up at 90 degrees and extend to 180 degree angle> and turn left <Head tilts left and left hand points to the left>

Follow the hallway and turn right <Head tilts right and right hand points right> then left <Head tilts left and left hand points left>

The location will be to your right <Head tilts right and right hand points right>

Right Behind: Olin 107

G1: No Gesture

Go to my right

When you reach the stairs turn right and go straight

Go through the double doors

Go through the double doors on your right

Follow the hallway and the location will be on your right

G2: Body rotation/posture

Go to my right <Turn body slight right and head tilts right>

When you reach the stairs turn right and go straight <Turn body slight right>

Go through the double doors

Go through the double doors on your right <Head tilts right>

Follow the hallway and the location will be on your right <Turn body slight right>

G3: Body + Arms

Go to my right <Turn body slight right, head looks right, right arm extends, and right hand is pointing to the right>

When you reach the stairs turn right and go straight <Head tilts right and right hand is point up at 90 degrees and extends to 180 degrees>

Go through the double doors

Go through the double doors on your right <Turn body slight right, head looks right, and right hand is pointing to the right>

Follow the hallway and the location will be on your right <Right hand is pointing to the right>

Left Up: Chemistry Stockroom S111

G1: No Gesture

Go to my right

When you reach the stairs to your right go up them

At the top of the stairs go straight and then through the double doors

Follow the hallway to the left and the location will be to your right (before the next set of double doors)

G2: Body rotation/posture

Go to my right <Turn body slight right and head tilts right>

When you reach the stairs to your right go up them <Turn body slight right and head tilts up>

At the top of the stairs go straight and then through the double doors

Follow the hallway to the left <Turn body slight left> and the location will be to your right (before the next set of double doors) <Head tilts right>

G3: Body + Arms

Go to my right <Turn body slight right, head looks right, right arm extends, and right hand is pointing to the right>

When you reach the stairs to your right go up them <Head tilts right and right hand is pointing up>

At the top of the stairs go straight and then through the double doors

Follow the hallway to the left <Head tilts left and left hand is pointing left> and the location will be to your right (before the next set of double doors) <Body turns slight right and right hand points right>

Right Up: Wold 128

G1: No Gesture

Go to my right

When you reach the stairs to your right go up them

There is a walkway on your right, follow it

Turn slight left and the location will be on your right

G2: Body rotation/posture

Go to my right <Turn body slight right and head tilts right>

When you reach the stairs to your right go up them <Turn body slight right and head tilts up>

There is a walkway on your right, follow it <Turn body slight right>

Turn slight left <Turn body slight left> and the location will be on your right <Head tilts right>

G3: Body + Arms

Go to my right <Turn body slight right, head looks right, right arm extends, and right hand is pointing to the right>

When you reach the stairs to your right go up them <Head tilts right and right hand is pointing up>

There is a walkway on your right, follow it <Body turns slight right, head tilts right, and right hand is pointing right>

Turn slight left <Head tilts left and left hand is pointing left> and the location will be on your right <Body turns slight right and right hand is point right.>

Fill out the appropriate part of the survey depending on what portion of the study you are in. Ratings range from 1 to 10 where 1 is a low rating and 10 is a high rating.

Location 1

How helpful were the instructions from the virtual agent?

1 2 3 4 5 6 7 8 9 10

Not Helpful Very Helpful

How natural were they?

1 2 3 4 5 6 7 8 9 10

Unnatural Very Natural

Location 2

How helpful were the instructions from the virtual agent?

1 2 3 4 5 6 7 8 9 10

Not Helpful Very Helpful

How natural were they?

1 2 3 4 5 6 7 8 9 10

Unnatural Very Natural

Location 3

How helpful were the instructions from the virtual agent?

1 2 3 4 5 6 7 8 9 10

Not Helpful Very Helpful

How natural were they?

1 2 3 4 5 6 7 8 9 10

Unnatural Very Natural

Location 4

How helpful were the instructions from the virtual agent?

1 2 3 4 5 6 7 8 9 10

Not Helpful Very Helpful

How natural were they?

1	2	3	4	5	6	7	8	9	10
Unnatural									Very Natural

Location 5

How helpful were the instructions from the virtual agent?

1	2	3	4	5	6	7	8	9	10
Not Helpful									Very Helpful

How natural were they?

1	2	3	4	5	6	7	8	9	10
Unnatural									Very Natural

Location 6

How helpful were the instructions from the virtual agent?

1	2	3	4	5	6	7	8	9	10
Not Helpful									Very Helpful

How natural were they?

1	2	3	4	5	6	7	8	9	10
Unnatural									Very Natural

What did you think the agent did that was helpful in the direction-giving?

What did you think the agent did that was confusing in the direction-giving?

Is English your native language?

Gender: **Age:** **Major/Occupation:**