Louwerse, M.M., Benesh, N., Hoque, M. E., Jeuniaux, P., Lewis, G. , Wu, J., & Zirnstein, M. (in press). Multimodal communication in face-to-face conversations. *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp.xx-xx). Mahwah, NJ: Erlbaum.

Multimodal Communication in Face-to-Face Computer-Mediated Conversations

Max M. Louwerse (mlouwerse@ memphis.edu)[a]
Nick Benesh (nbenesh@memphis.edu)[a]
Mohammed E. Hoque (mhoque@memphis.edu)[b]
Patrick Jeuniaux (pjeuniau@memphis.edu)[a]
Gwyneth Lewis (glewis1@ memphis.edu)[a]
Jie Wu (jwu4@ memphis.edu)[c]
Megan Zirnstein (mzirnstn@ memphis.edu)[a]

Department of Psychology / Institute for Intelligent Systems[a]
Department of Electrical and Computer Engineering / Institute for Intelligent Systems[b]
Department of Computer Science / Institute for Intelligent Systems[c]
Memphis, TN 38152 USA

Abstract

Multimodal communication involves multiple communicative channels including speech, facial movement, and gesture. Relatively few studies on how various communicative modalities are aligned in natural, face-to-face communication exist. As part of a larger project, the current study investigates how discourse structure, speech features, eye gaze, and facial movements interrelate during a map coordination task. The study thereby sheds light on multimodal communication in humans and gives guidelines for the development of embodied conversational agents.

Keywords: multimodal communication; discourse structure; dialogue; embodied conversational agents

Introduction

Multimodal communication is comprised of various modalities such as speech, facial movement, and gesture. Avoiding gestures while talking on the phone is difficult, not looking at someone in a face-to-face setting is challenging, , and listening to someone in a face-to-face setting without acknowledging information with the occasional *mhmm*'s and head nods, again, is a challenge. Despite the deceptively simple appearance of these communicative tools, little is understood regarding their interaction and alignment. Modalities such as speech, discourse, facial movements, eye gaze, and gesture seem to be intrinsically related. At the same time, little is known about how they interrelate and how they are aligned.

Knowing the nature of the interaction of modalities and their alignment can shed light on various areas of cognitive science. From a psychological perspective, an understanding of the interplay of modalities can help us understand language and communication (Clark, 1996). Limited experimental research is available that can help determine whether modalities can be substituted or whether they are complementary (cf. Doherty-Sneddon, et al., 1997).

From an educational perspective, an understanding of modalities can help answer questions regarding student motivation, interest, and confusion, as well as how instructors and tutors can monitor and respond to these cognitive states (Kort, Reilly & Picard, 2001). But with little information available concerning the conditions under which students use facial movements or eye gaze, tapping into students' cognitive states is difficult (cf. Graesser, et al., in press).

From a computational perspective, an understanding of the interplay between modalities can help in the development of animated conversational agents (Louwerse, Graesser, Lu & Mitchell, 2005). These agents maximize the availability of both linguistic (semantics, syntax) and paralinguistic (pragmatic, sociological) features (Cassell & Thórisson, 1999; Massaro & Cohen, 1994; Picard, 1997). But without experimental data on multimodal communication, the guidelines for implementing human-like multimodal behavior in agents are missing (cf. Cassell, et al., 1994).

The current paper presents some initial results of an extensive data collection study of 256 conversations from 64 participants, all native speakers of English, and totaling 35 hours. Conversations were monitored for dialogue acts, speech, facial movements, gesture, eye gaze, and route drawing accuracy. The data from this study will shed light on human multimodal communication and will provide guidelines for the development of natural, embodied conversational agents.

Map Task Dialogues

Though there is considerable amount of research on multimodal communication (Argyle & Cook, 1976; Doherty-Sneddon, et al. 1997; Ekman, 1979; Goldin-Meadow, 2003; Louwerse & Bangerter, 2005; McNeill, 1992), this research can be characterized by the fact that 1) pairs of modalities are considered, so that it remains unclear whether multiple modalities are mutually substitutable; 2) language situations are highly diverse, making it difficult to interpret why certain multimodal behavior occurs; and 3) dialogue is unpredictable, making it hard to model when modalities behave in certain ways.

In the current research project on multimodal communication in humans and agents (Louwerse, et al., 2004), we are investigating the interaction between dialogue act, speech, eye gaze, facial movements, gesture, and map drawing. The project aims to determine how these modalities are aligned, whether, and if so when, these modalities are observed, and whether the correct use of these channels actually aids comprehension.

Due to the inherent complexity of multimodal communication, controlling for genre, topic, and goals during unscripted dialogue is crucial. With these concerns in mind, we used the Map Task scenario (Anderson, et al., 1991), a restricted-domain, route-communication task. In the Map Task scenario it is possible for experimenters to determine exactly what each participant knows at any given time. In this scenario, the Instruction Giver (IG) coaches the Instruction Follower (IF) through a route on the map. By way of instruction, participants are told that they and their interlocutors have maps of the same location, but drawn by different explorers, and so are potentially different in detail. They are not told where or how the maps differ.

The present paper reports a preliminary analysis of the experimental data gathered for this project.

Method

For the current paper, we randomly sampled 16 of 256 available conversations, totaling 72 minutes of dialogue with different participants and different maps for each conversation. Data from each conversation consisted of recordings of participants' facial movements, gestures, speech, eye gaze patterns (for IG), and map drawings (for IF). Each of the participants performed the role of IG (4 conversations in a row) and the role of IF (4 conversations in a row). In each conversation, different maps were used that varied in terms of homogeneity of objects. An example of maps for the IG and IF is given in Figure 1.

Participants

Of the 64 participants in the total data set, 32 were included in this analysis yielding a representative sample. Of these participants, 21 were female. Sixteen participants were Caucasian, 14 African-American, and 2 Asian-American, all being native speakers of English.

Materials

Sixteen different maps were used, each varying according to the presentation of landmarks, route shape, and method of distortion in the IF map. For instance, IF's maps were distorted with blurred out portions of the map. Four possible route shapes were used. As for landmark presentation, maps either had mixed or single themes. For example, a mixed landmark map would include a majority of landmarks from a theme (e.g., birds) and a few randomly selected landmarks from other themes (e.g., aliens), as in Figure 1.

Apparatus

Participants' communication was recorded using five Panasonic camcorders, two capturing the faces of each dialogue participant (PV-GS31), two capturing the upper torsos of each participant (PV-GS150), and one capturing both participants from a bird's eye view (PV-GS150). Eye gaze was recorded for the IG only using an SMI iView RED

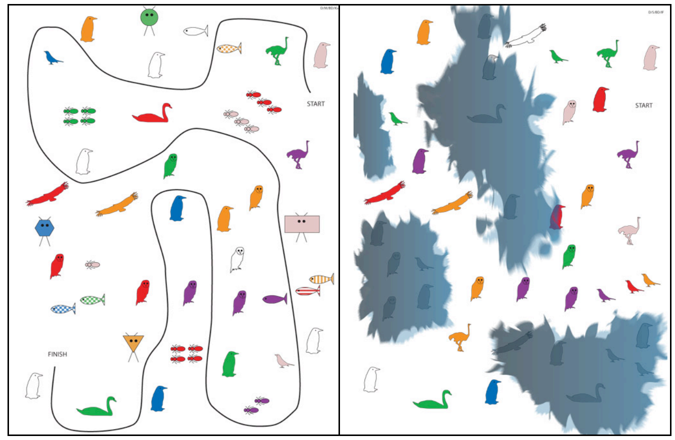

Figure 1. Examples maps for the IG (left) and the IF (right)

remote eye tracker. Speech was recorded using a Marantz PMD670 recorder, whereby IG and IF were recorded on two separate (left and right) channels using two AKG C420 headset microphones. Two high-resolution webcams were used for the interface. The IF drawings of the routes on the screen were recorded both spatially and temporally.

Procedure

Participants were seated in front of each other but were separated by a divider to ensure that they could not see each other. They communicated through microphones and headphones, and could see the upper torso of their dialogue partner through the webcam and the map on a computer monitor in front of them. This computer-mediated session, using webcams, was necessary for eye tracking calibration, as well as to reduce torso movement.

The IG was presented with a colored map with a route (similar to the one presented in Figure 1). The IG's task was to communicate the route to the IF as accurately as possible. To ensure on-task conversations, participants were promised extra payment if the route drawn by the IF matched the route on the IG's map.

Equipment was calibrated before the start of each conversation. The five camcorders were positioned and focused in order to best capture the facial movements and upper torso movements of each participant. The eye gaze of the IG was calibrated using nine calibration points on the computer monitor. To avoid interruption of the dialogue, calibration only occurred once per map. When calibration was lost during the study, recording of eye tracking data was discontinued and this part of the data was eliminated from the analysis. Each conversation started with a flash of light and the sounding of a brief tone, in order to ensure alignment of the different channels in later analysis.

Results and Discussion

Dialogue acts

The 12 dialogue acts (DA) that are typically used for Map Task coding were used (Carletta et al., 1997; Louwerse &

Crossley, 2006). A description and an example of these DAs is presented in Table 1. Two trained coders manually coded the utterances of half of the conversations as one of the twelve DAs. Inter-rater reliability between the coders in terms of Cohen's Kappa was at .67. Coders resolved the conflicts, primarily relating to the acknowledgment DA and coded the remaining transcripts for DAs. An overview of the frequency of the DAs is presented in Table 1.

Speech

Speech features related to pitch, pause, and speaking rate were calculated for each dialogue act, using *Praat* speech processing software (Boersma & Weenink, 2006).

Pause was analyzed, using the upper intensity limit and minimum duration of silence. In measurement of intensity, minimum pitch specifies the minimum periodicity frequency in any signal. In our case, 75 Hz for minimum pitch yielded a sharp contour for the intensity. Audio segments with intensity values less than 58 dB and with a duration of silences longer than 0.2 seconds were classified as pauses. Features such as number of pauses, maximum duration of pauses, average duration of pauses, and total duration of pauses per dialogue act were computed.

Pitch information from each dialogue act speech segment was computed using the autocorrelation method processed through *Praat* speech processing software. The lower and upper thresholds for pitch were set to 75Hz and 400Hz, respectively. In other words, only pitch components ranging from 75Hz to 400Hz were considered for our analysis. Speaking rate was computed as the ratio of voiced frames (1/voiced frames).

Of course, the maximum threshold for intensity is dependent on individuals' vocal tone, gender, and other issues. Also, speaking rate is different for each individual. Our future efforts will therefore include using an adaptive approach to automatically predict the maximum threshold for intensity and minimum duration for silence, per participant.

Facial movements

Standard emotion coding schemes like Ekman, Friesen, Wallace, and Hager's (2002) Facial Action Coding Scheme (FACS) are problematic for Map Task scenarios because emotions like disgust, anger, and sadness do not occur frequently during these interactions. Implementing a subset of the action units themselves is, however, beneficial. A total of 20 facial movements were coded for. Their labels, descriptions, and frequencies are presented in Table 2.

Three coders rated four conversations for facial movement. Cohen's Kappa was .77 for the head, .73 for eye brows, .79 for the eyes, and .85 for the mouth, yielding an overall inter-rater reliability for facial movement coding of .78. Coders independently rated the remaining conversations. Facial movements were first coded then aligned with speech in order to analyze their interaction over time.

Eye gaze

Eye fixation for the IG only was recorded in order to implement findings in the embodied conversational agent fulfilling the role of IG in the project. Two main areas of interest were identified: the IF's webcam image and the IG's map. In addition, eye fixation on the rest of the screen, fixation off of the screen, and lost fixation times were recorded. 74.1% of the conversation had fixations recorded on the screen. Of these fixations, 23.5% were not aimed at the map or the IF, 9.7% was aimed directly at the IF, and 66.8%, on the map.

Map drawings

The route drawn by the IF is useful to determine the extent to which the IF deviated from the route on the IG's map. This gives us an idea as to whether communication between the IG and IF was executed successfully. To obtain this information, we computed the minimal difference between the given route on the IG's map and the drawn route on the IF's map. The average difference between the ideal route on the IG's map and the drawn IF's map was 11.53 pixels (*SD* = 20.47), which translates to .58 cm, with a min. difference of 0 and a max. difference of 134 (approximately 6.8 cm).

Table 1. The 12 DAs used in the Map Task, a description, an example, and frequencies for the selected conversations

Dialogue Act	Description	Example	IG	IF
INSTRUCT	Commands partner to carry out action	*Go down between the blue and the red car.*	698	12
EXPLAIN	States information not directly elicited by partner	*Ok I went the wrong way.*	234	76
CHECK	Requests partner to confirm information	*So, between the black and the grey one?*	10	56
ALIGN	Checks attention, readiness, agreement of partner	*Ok, do you see those two blue cars?*	38	1
QUERY-YN	Yes/no question that is not CHECK or ALIGN	*Do you see that?*	218	50
QUERY-W	Any query not covered by the other categories	*If I'm at the red car what do I do there?*	22	23
ACKNOWL	Verbal response minimally showing understanding	*Uh huh.*	176	470
REPLY-Y	Reply to any yes/no query with yes-response	*Yeah, start at the top.*	59	135
REPLY-N	Reply to any yes/no query with no-response	*No,go like above the puddle.*	8	6
REPLY-W	Reply to any type of query other than 'yes or 'no'	*It goes below.*	17	20
CLARIFY	Reply to question over and above what was asked	*So you'll be between the blue and red car.*	21	41
READY	Preparing conversation for new dialogue game	*Alright. We're going to move to the left.*	22	6
UNCODBL			20	4

Table 2. Facial movements used in the coding, a description and their average frequencies by IG and IF
Note: AU = Corresponding Action Unit, IG = Information Giver, IF = Information Follower

	Label	Description	AU	IG	IF
Head	Forward movement	Slow forward head movement towards the monitor	AU57	46.38	34.06
	Backward movement	Slow backward head movement away from the monitor	AU58	44.00	28.13
	Left tilt	Head tilting resulting in head cocked to left	AU55	21.19	14.88
	Right tilt	Head tilting resulting in head cocked to the right	AU56	17.06	14.25
	Nodding fast	Quick downward and upward movement of head	n/a	16.13	13.06
	Nodding slow	Slow downward and upward movement of head	n/a	6.19	7.06
	Turning (left)	Neck movement results in face orienting towards the left	AU51	22.06	19.06
	Shaking right	Neck movement results in face orienting towards the right	AU52	22.25	19.44
Eye / eyebrows	Brows (up)	Inner or outer portion of the eyebrows are pulled upwards.	AU1/2	10.38	6.31
	Brow (down)	Lowering of eyebrows as in frowning	AU4	0	0
	Asymmetrical	Only one eyebrow (left or right) is raised.	n/a	1.81	1.13
	Rapid blinking	Eyes close and open very quickly with no pause in succession	AU45	0.69	1.13
	Squinting	Narrowed eye aperture.	AU44	4.56	2.56
	Widening	Widens the eye aperture	AU5	1.00	0.44
	Rolling	Upward rolling of the eyes.	M68	0	0
Mouth	Smile	Pulls the corners of the lips back and upward	AU12	2.00	2.56
	Lip tightener	Tightens the lips, making lips appear more narrow	AU23	2.00	4.75
	Lip pucker	Pushes the lips of the mouth forward and pulls medially	AU18	0.00	0.44
	Mouth open	Mouth is in O shape	AU25/26	0.88	2.63
	Biting lip	Teeth biting the lip (teeth may or may not be viewable)	AD32	1.13	3

Correlations between modalities

While the correlational data are at best exploratory in nature, they can provide the first steps in helping to resolve the complex issue of how to associate these communicative modalities with one another. The data for all modalities were time aligned on 250 millisecond intervals. Though eye gaze, map drawings, and speech features allowed for smaller time frames (< .016 msecs), facial movement coding did not. We therefore aggregated the data to a unit of analysis of 1 second. To avoid a Type I error, only those correlations were considered that were significant at p <.001.

Correlations were found between route difference, facial movements, and speech features. Larger route differences correlated with a smile in the IF (r = .13) and a head tilt to the left (r = .46). These facial movements may signify surprise or confusion (Ekman, 1979). For instance, asymmetrical eye brows in the IF is associated with squinting of the eyes in the IG (r = .84) as well as a shaking of the head of the IG (r = .68). But correlations were also found within a speaker. For instance, the IF's eye brows going up correlated with the squinting of the eyes (r = .62) as well as widening of the eyes (r = .79) in lip biting (r = .64) and smiling (r = .77) in the IF.

Dialogue partners seem to copy each other's modalities. When the IG nodded slowly, the IF nodded slowly (r = .87), and smiles from the IG were accompanied with a smile from the IF (r = .67).

Correlational data like these show that different communicative channels are correlated with other channels, both within a speaker and across dialogue partners. They thereby confirm the complexity of the interrelations between modalities and make it hard to interpret the results. In addition, these correlations make it difficult to provide guidelines for agent development. Therefore, what is desirable is one modality that can be used as a basis onto which other modalities can be mapped.

Dialogue structure and its relations to modalities

One of the aims of the current project is, as described, the development of an animated conversational agent that can interact with a human dialogue partner and behave similarly to the human dialogue partner in terms of using modalities like dialogue structure, speech features, eye gaze, and facial movements.

The correlations presented above may give insights into the interrelations of the various modalities. They are of limited use in agent development, however, because we lack a basis onto which the modalities can be mapped. In order to solve this problem, we used dialogue structure as the basis of comparison, because it provides precise cues regarding the message's meaning and the speaker's meaning.

Using dialogue structure has an additional advantage, in that algorithms have been put in place to classify utterances into dialogue acts (Louwerse & Crossley, 2006). This means that once the agent has identified a dialogue act, it can elicit the modalities that are most affected by this dialogue act. The agent can then identify the dialogue act in its own utterance and elicit the relevant modalities, but also identify the dialogue act in the IF utterance and respond with the appropriate modalities.

Multiple regression analysis allows us to infer how accurately predictions can be made of a particular modality if the dialogue act is known. We conducted a multiple regression analysis with the 12 dialogue acts as dummy-coded independent variables entered in a Stepwise fashion with the modalities as dependent variables. Scores for the dependent variables were averaged by the IG and IF dialogue acts, resulting in 279 cases per modality. We will focus only on those findings wherein the overall dialogue structure explained a significant amount of the variance ($R^2 > .15$, $p < .01$) and where the standardized regression coefficients (β) for individual dialogue acts explained a significant ($p < .001$) amount of the variance.

The results, summarized in Table 3, show which modality's variance can be significantly explained per dialogue act per participant role (IG/IF). For instance, when the IG uses a REPLY-N dialogue act, typically a no answer to a yes/no question, it is likely that the IG's eyebrows will be up, that the IG is expressing a smile, and that the speaking rate will be high. While the REPLY-N is expressed, the IF tends to perform a backward head movement. Similarly, when the IF asks a QUERY-W, typically a wh-question, there tends to be an increase in rising and falling pitch in speech, upward eyebrow movements, and rapid blinking in the IF, while the IG's eyes are widening.

In terms of the development of human-like, embodied conversational agents, results like these can provide guidelines concerning when to use what modalities. Therefore, if the system has identified a DA in the IF, it then knows which modalities to activate.

Conclusion

Even though multimodal communication includes perhaps the most fundamental forms of communication, there are relatively few studies on how various communicative modalities are aligned in natural, face-to-face communication. The reasons for this are simple. Collecting data, whereby a range of modalities are recorded properly, all participants elicit in natural dialogue while being recorded, and all data can be temporally aligned, is difficult. Moreover, when such naturalistic data is collected, researchers are confronted with a wealth of possible variables.

The present study, part of a larger project, provides insight into how eye gaze, facial movements, speech features, map drawings, and dialogue structures correlate with each other and which dialogue acts best predict the expression of a particular modality. We believe that by incorporating the findings of human-to-human, multimodal communication into an artificial agent, the agent will be able to interact with humans more naturally and fluently. Based on the sample of data discussed in this paper, we have provided preliminary guidelines for the development of embodied, conversational agents and shown that multiple communicative channels are really interdependent communicative channels.

Acknowledgments

This research was supported by grant NSF-IIS-0416128. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding institution. We would like to thank Ellen Bard and Markus Guhe for the creation of the maps and their help in setting up the study.

Table 3. Regression relationships between dialogue act and modalities. For all relations included $\beta > .31$, $p < .001$
Note: All relations are positive: presence of a dialogue act increases chance of presence of modality

IG	**IG**	**IF**
ACKNOWL	speaking rate	
ALIGN	speaking rate	
CHECK	squinting eyes	
EXPLAIN	freq. falling and rising pitch, average duration pauses, freq. pauses	
INSTRUCT	freq. falling and rising pitch, average duration pauses, freq. pauses	
QUERY W	freq. falling and rising pitch	
QUERY YN	freq. falling and rising pitch, average duration pauses, speaking rate	
READY	head backward movement, speaking rate	
REPLY N	eyebrows up, smile, speaking rate	head backward movement
REPLY Y	speaking rate	nodding slow
UNCODABLE	laughing	
IF	**IF**	**IG**
ACKNOWL.	freq. falling and rising pitch, freq. pauses, speaking rate	
ALIGN	nodding fast, rapid blinking, speaking rate	turning right
CHECK	freq. falling and rising pitch, freq. pauses	
CLARIFY	freq. falling and rising pitch	
EXPLAIN	freq. falling and rising pitch, freq. pauses, speaking rate, dur. Pauses	nodding slow
INSTRUCT	freq. rising pitch, eyebrows asymmetrical	
QUERY W	freq. falling and rising pitch, eyebrows up, rapid blinking	eyes widening
QUERY YN	freq. falling and rising pitch, freq. pauses	
READY	speaking rate, left tilt	lip tightener, head tilt right,
REPLY_N	mouth pucker, mouth smile, turning right	mouth open
REPLY W	freq. falling and rising pitch, freq. pauses, route distance, left tilt, dur. pauses	
REPLY Y	biting lip, lip tightener, speaking rate	squinting eyes

References

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech*, *34*, 351-366.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Boersma, P., & Weenink, D. (2006). *Praat: Doing phonetics by computer* (Version 4.4.06) [Computer program]. Retrieved January 30, 2006, from http://www.praat.org/

Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, *23*, 13-31.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., & Stone, M. (1994) Animated Conversation: Rule-Based Generation of Facial Expression, Gesture and Spoken Intonation for Multiple Conversational Agents. *Proceedings of SIGGRAPH '94*, 413-420.

Cassell, J., & Thórisson, K. R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, *13*, 519-538.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Doherty-Sneddon, G., Anderson, A. H., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*, *3*, 105-125.

Ekman, P. (1979). About brows: Emotional and conversational signals. In M. von Cranach, K. Froppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: Claims and limits of a new discipline: Contributions to the colloquium* (pp. 169-248). Cambridge: Cambridge University Press.

Ekman, P., Friesen, Wallace V., & Hager, J.C. (2002). *Facial Action Coding System (FACS)*. CD-ROM.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Graesser, A. C., D'Mello, S. K., Craig, S. D., Witherspoon, A., Sullins J., McDaniel B., & Gholson, B. (in press). The relationship between affective states and dialog patterns during interactions with AutoTutor. *Journal of Interactive Learning Research.*

Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In *Proceedings of the International Conference on Advanced Learning Technologies (ICALT 2001)*, Madison Wisconsin, August 2001.

Louwerse, M. M., & Bangerter, A. (2005). Focusing attention with deictic gestures and linguistic expressions. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society.*

Louwerse, M. M., Bard, E.G., Steedman, M., Hu, X., & Graesser, A. C. (2004). *Tracking multimodal communication in humans and agents*. Technical report, Institute for Intelligent Systems, University of Memphis, Memphis, TN.

Louwerse, M. M., & Crossley, S. (2006). Dialog act classification using n-gram algorithms. In *Proceedings of the Florida Artificial Intelligence Research Society International Conference (FLAIRS).* Menlo Park, CA: AAAI Press.

Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology*, *19*, 1-12.

Massaro, D. W., & Cohen, M. M. (1994). Visual, orthographic, phonological, and lexical influences in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1107- 1128.

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

Picard, R. (1997). *Affective computing*. Cambridge, MA: MIT Press.