

Facial Expressions and Politeness Effect in Foreign Language Training System

Ning Wang¹, W. Lewis Johnson², Jonathan Gratch¹

¹USC Institute for Creative Technologies
13274 Fiji Way, Marina del Rey, CA 90202 USA
{nwang,gratch}@ict.usc.edu

²Alelo Inc.
12910 Culver Boulevard Suite J, Los Angeles, California 90066 USA
ljohnson@alelo.com

Abstract. Previous studies on the Politeness Effect show that using politeness strategies in tutorial feedback can have a positive impact on learning (McLaren et al. 2010; Wang and Johnson 2008; Wang et al. 2005). While prior research efforts tried to uncover the mechanism through which the politeness strategies impact the learner, the results were inconclusive. Further, it is unclear how the politeness strategies should adapt over time. In this paper, we analyze the video tapes of participants' facial expression while interacting with a polite or direct tutor in a foreign language training system. The Facial Action Coding System was then used to analyze the facial expressions. Results show that as social distance decreases over time, polite feedback is received less favorably while the preference for direct feedback increases.

Keywords: politeness effect, facial expression, facial action coding system, second language acquisition

1 Introduction

In recent years, there has been rigorous research on pedagogical agents' ability to facilitate learning (Atkinson, 2002; Johnson et al. 1998; Lester et al. 2000; Moreno, 2005). While some research focused on the agent's appearance and voice (Baylor, 2005; Baylor et al. 2003; Graesser et al. 2003; Moreno and Mayer, 2000; Moreno et al. 2001), we focused instead on the way agent's feedback is delivered. We conducted a series of studies on the use of politeness strategies in tutorial feedback and showed that the pedagogical agent's use of politeness strategies can promote better learning results (Wang et al. 2005; Wang and Johnson, 2008). This *politeness effect* was later tested in real classroom settings (McLaren et al. 2007). The latest study shows that individual differences, such as level of domain knowledge, can impact the *politeness effect* (McLaren et al. 2010). While the *politeness effect* was well studied in terms of its impact on learning, it was unclear what may be the mediating factors. In our earlier

analysis, we hypothesized that motivation, in particular self-efficacy and sense of autonomy, are the factors through which politeness operate upon (Wang and Johnson 2008; Wang et al. 2005). However, results from the analysis were inconclusive.

Brown and Levinson (1987) argue that people in all cultures have face wants. The notion of face wants refers of two specific kinds of desires: the desire to be unimpeded in one's action (negative face), and the desire to be approved of (positive face). The use of politeness strategies is to mitigate the threat to face wants and facilitate harmonious interaction. An alternative explanation for the *politeness effect* could simply be that the use of politeness strategies puts the learner in an affective state that is more suitable for learning. Research on emotion and emotional expression shows that people categorize facial expressions of emotions in a similar way across cultures, and that people produce simulations of facial expressions that are characteristic of each specific emotion (Ekman, 1993). In our study of the politeness effect in a foreign language culture training system, we recorded participants' facial expressions while they interacted with the system. In this paper, we present our investigation of learners' affective states through analysis of learners' facial expressions.

Another question left unanswered is how adaptive the politeness strategies are over time, when used in tutorial feedback. The proper level of politeness depends on the potential threat of a communicative act. In the Brown and Levinson model (1987), evaluation of face threat depends upon several factors. First, the relative weight of different face threats is culturally dependent. This culture dependency is defined as the ranking of impositions by the degree to which they are considered to be interfering with one's want of autonomy and approval. Second, the weight of a face-threatening act also depends upon the relative power between the speaker and the listener. Tutors generally have power relative to learners, so we would generally expect tutors to make use of weaker politeness strategies when speaking to learners than the learners use in reverse. Finally, the weightiness of a face threat depends upon the social distance between the two parties. As two people interact over time, their social distance often decreases, reducing the severity of face threatening acts and increasing the likelihood of actions such as direct requests that lack face-saving features. In tutoring sessions, the first two factors, culture and relative power, do not change much over time. However, the social distance between the learner and the tutor could decrease. If the politeness strategies do not adjust to the change of social distance over time, would the learner react to the feedback differently?

In this paper, we investigate the following research hypotheses:

- H1. Learner affect is a mediating factor between politeness and learning.
- H2. The use of politeness strategies in tutorial feedback needs to adapt to the change in social distance between the learner and pedagogical agent over time.

2 Facial Action Coding System

To analyze the facial expressions, we used the Facial Action Coding System (FACS) (Ekman and Friesen, 1978). The FACS is arguably the most widely used method for coding facial expressions in the behavioral sciences. The system describes facial expressions in terms of 46 component movements, which roughly correspond to the individual facial muscle movements. FACS provides an objective and comprehensive way to analyze expressions into elementary components. Because it is comprehensive, FACS has proven useful for discovering facial movements that are indicative of cognitive and affective states (Ekman and Rosenberg, 2005).



Fig. 1. From left to right, pictures of facial display of AU 4 (Brow Lower), AU 9 (Nose Wrinkle), AU 10 (Upper Lip Raise) and AU 12 (Lip Corner Puller).

3 CERT

The primary limitation to the widespread use of FACS (Ekman and Friesen, 1978) is the time required to code. FACS was developed for coding by hand, using human experts. It takes over 100 hours of training to become proficient in FACS, and it takes approximately 2 hours for human experts to code each minute of video.

Table 1. Action Units automatically coded by CERT

<i>Action Unit</i>	<i>Description</i>	<i>Action Unit</i>	<i>Description</i>
1	Inner Brow Raise	15	Lip Corner Depressor
2	Outer Brow Raise	17	Chin Raiser
4	Brow Lowerer	18	Lip Pucker
5	Upper Lid Raise	20	Lip Stretch
6	Cheek Raise	23	Lip Tightener
7	Lids Tight	24	Lip Presser
9	Nose Wrinkle	25	Lips Part
10	Upper Lip Raiser	26	Jaw Drop
12	Lip Corner Puller	27	Mouth Stretch
14	Dimpler	28	Lips Suck

To analyze the facial expressions more efficiently, we processed our video data through the Computer Expression Recognition Toolbox (CERT) developed by the University of California at San Diego (Bartlett et al. 2004). CERT is a user independent fully automatic system for real time recognition of facial actions from the Facial Action Coding System (FACS). The current version of CERT produces a 20 channel output stream. Each output stream channel consists of one real valued number for an Action Unit (AU), for each frame of the video. The real valued number indicates the distance to the separating hyper-plane for each classifier Support Vector Machine classifier. Previous work showed that the distance to the separating hyper-plane (the margin) contained information about Action Unit intensity (Bartlett et al. 2006). The 20 Action Units CERT output are shown in Table 1. Previous work (Susskind et al. 2007) shows that CERT performs comparably to human observers in the discrimination of distinct basic emotion classes and judgments of the similarity between distinct basic emotions.

In the investigator's guide to FACS, Ekman and Friesen (1978) describe the Action Units that are generally associated with facial expressions of different emotions. For example, facial expressions of joy typically include the activation of AU 12 (Lip Corner Puller) and AU 6 (Cheek Raise). AU 9 (Nose Wrinkle) or AU 10 (Upper Lip Raise) is often seen in facial expressions of disgust. Following the investigator's guide, we used AU 6 and AU 12 as indications of positive emotional facial expressions and AU 4, AU 9 and AU 10 as indications of negative emotional facial expressions (Figure 1). Positive and negative emotional facial expressions can certainly include other Action Units. However, from the actions units that can be automatically detected by CERT so far, these are the most commonly associated with positive and negative emotional facial expression.

In the analysis of learner facial expressions when interacting with the AutoTutor, McDaniel et al. (2007) correlated the learner reported affective states and FACS coding from two independent coders. The analysis identified eight Action Units (AU1, AU4, AU7, AU12, AU25, AU26, AU43 Eye Closure and AU45 Blink) that significantly correlated with five affective states (Boredom, Confusion, Delight, Frustration and Neutral). In this paper, we focus on analyzing facial expressions indicated by six of these eight Action Units (excluding AU 43 AND 45 since CERT does not output these two at the moment) and the ones generally associated with positive and negative emotions as described above.

4 Data Description

Tactical Iraqi is one of several game-based courses developed by Alelo Inc. It is a training system that supports individualized language learning and helps military service members quickly acquire functional communication skills. Tactical Iraqi includes three modules: the Skill Builder, the Mission Game and the Arcade Game. The Skill Builder consists of interactive lessons and exercises, and interactive game experiences. Learners use headset microphones to interact with the software, along with a keyboard and mouse. Lessons, exercises, and game experiences all involve

speaking in the target language; speech recognition software is used to interpret the learner's speech. The current study focuses on Skill Builder only. More information on the Arcade Game and Mission Game can be found in Johnson (2007).

To investigate the effect of politeness strategies in tutorial feedback, we created two types of feedbacks: polite feedback which is phrased using various politeness strategies and direct feedback which is phrased without any politeness strategies. An example of direct feedback is "No, that means 'This is a sergeant.' Try again." An example of polite feedback is "It's usually hard to get answers to this question right, but that means 'This is a sergeant.' How about we try it again?" Details about the politeness strategy can be found in Wang and Johnson (2008).

Sixty-one volunteers (59% women, 41% men) from the greater Los Angeles area participated in the study. They were recruited by responding to recruitment posters on Craigslist.com and were compensated \$40 for three hours of their participation. On average, the participants were 38.4 years old (min=21, max=63, std=11.5). The study design was a between-subjects experiment with two conditions: Polite (n=31) and Direct (n=30), to which participants were randomly assigned.

Participants filled out the pre-questionnaire packet and started training in the Skill Builder in Tactical Iraqi. Participants in the Polite condition received polite feedback while participants in the Direct condition received direct feedback. Participants completed one hour training in day 1, returned to the lab next day and completed another hour of training. At the end of their training in day 2, participants were asked to write down the name of the lessons they took in Skill Builder. Then participants filled out the post-questionnaire packets and took the quizzes from the lessons they took in Skill Builder. The quizzes were constructed by our research team.

Learning Gains were measured using quizzes at the end of each lesson in the Skill Builder. The quizzes contain three types of questions. First type of question is Utterance-Formation questions, where participants answer questions by recording their own speech. The second type is Multiple-Choice questions. The third type is Match-Item questions, where participants match phrases in Iraqi Arabic to translations in English. Each correct answer gets 1 point. Participants took quizzes from all the lessons that they took during the 2 hour training.

Two indexes of motivation were measured: self-efficacy and perceived autonomy. Self-efficacy was measured both in the pre-training questionnaire ($\alpha=.829$) and the post-training questionnaire ($\alpha=.713$). Items from the self-efficacy scale are modified from the scales published in Boekaerts (2002). The difference between pre and post training results allows interpretation of how self-efficacy changes due to the training. Sense of autonomy ($\alpha=.885$) was measured only in the post-training questionnaire. The measure was designed by our research team. Example items from the autonomy measure include "I feel the system was deciding what I should do next for me."

5 Results

Data from eleven sessions were excluded. Two sessions were excluded because a computer crash and a speech recognizer malfunction. One session was excluded

because a participant's hearing and speech impairment. Four sessions were excluded because the participants "cheated" on the post-test. Four other sessions were excluded because CERT failed to locate the participant's face in the video, which is a pre-step to facial expression coding. As a result, data from 46 sessions ($N_{\text{Polite}} = 22$, $N_{\text{Direct}} = 24$) were included in the analysis. In this paper, we focus on the analysis of facial expressions. Results on learning and motivation are in Wang and Johnson (2008).

To process the CERT output, we adopted the statistical method Littlewort and her colleagues used to differentiate posed and genuine pain (Littlewort et al. 2007). This method strips out the individual variance in CERT output, e.g. different individuals have different baselines. It also sums up the overall activity of the Action Unit. We calculated the mean of the Z-scores for each participant (speaker only) and each AU detector as $Z=(x-\mu)/\sigma$, where (μ,σ) are the mean and variance for the output of the parts of each participant's video where the face was relatively neutral. Duration of the neutral face range from 3 seconds to 37 seconds (100 frames to 1114 frames).

Overall, we did not find any significant difference on individual Action Units between the Polite and Direct group. Correlation analyses showed that there was no significant correlation between the quiz score, self-efficacy and autonomy with any facial Action Units we tested. Previous analysis showed that politeness did not impact the overall quiz score but did help the learner perform better on more difficult and complex problems – the Utterance Formation quiz questions (Wang and Johnson, 2008). Further correlation analysis shows that AU 7 (Lids tight) is positively correlated with the Utterance Formation quiz score ($r=.315$, $p=.033$). We followed up with a stepwise linear regression using the Utterance Formation quiz score as the dependent variable, the experiment condition and the Action Units as independent variables. The model kept AU 7 and excluded the experiment condition and other Action Units. The resulting model with AU 7 is statistically significant ($F=4.835$, $p=.033$). Since previous study showed that age can significantly impact performance on the recall test (Wang and Gratch, 2009), we added age as an independent variable to this model. The resulting model with AU 7 and age is statistically significant ($F=5.193$, $p=.01$). This means that the learner's age and AU 7 activity are significant predictors of his/her performance on difficult and complex problems.

To investigate whether the learner perceived the politeness strategies of the same politeness level differently over time, we conducted a General Linear Model Repeated Measure analysis using activation of facial Action Units in the first session and second session (day 1 and day 2) as the dependent variable and the experiment condition as the independent variable. Results show that there is a significant interaction of AU 12 activity over time and experiment condition. ($p_{\text{Time}}=.743$, $p_{\text{Time*Condition}}=.041$). Figure 2 shows that activation of AU 12 decreases over time for learners in the Polite group. But for learners in the Direct group, their AU 12 activity increases from day 1 to day 2. AU 12 is strongly correlated with joy and delight (Ekman and Friesen, 1978, McDaniel 2007). This means that learners in the Polite condition initially enjoyed the polite feedback but found the feedback less enjoyable over time. On the other hand, learners in the Direct condition grew increasingly accustomed to the direct feedback and perceived it more favorably over time. We did not find any significant interaction of AU 6 activity over time and experiment

condition. However, the overall level of AU 6 activity is significantly correlated with AU 12 activity ($p=.003$, $r=.423$).

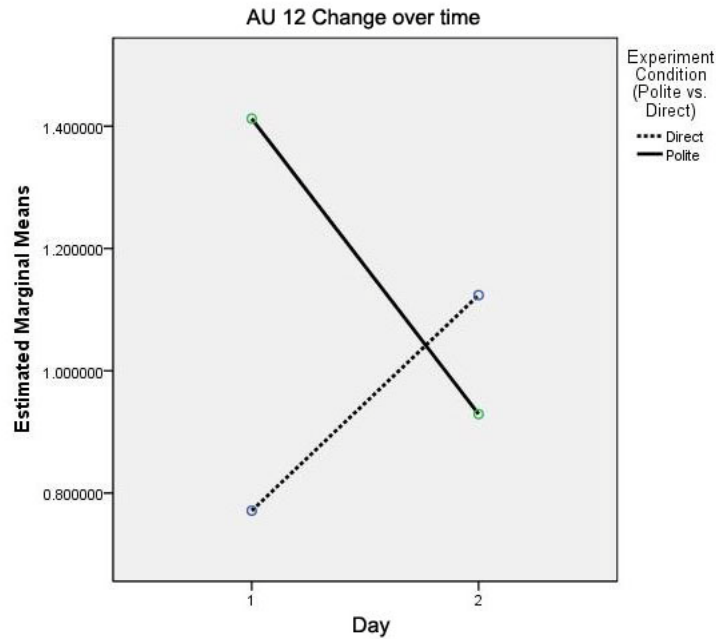


Fig. 2. Activity of AU12 changes differently from the first session (day 1) to second session (day 2) for learners in the Polite and Direct group.

6 Discussion

In this paper, we seek to test two hypotheses regarding the *politeness effect*. First, we hypothesize that learner affect could be a mediating factor between politeness and learning. This hypothesis was not supported. Results show that there was no significant difference on any facial Action Units between the polite and direct conditions. However, correlation analysis shows that AU 7 is significantly correlated with performance on difficult and complex problems. AU 7 is more predictive of learner performance than experiment manipulation. Previous studies showed that AU 7 is positively correlated with confusion and delight, and negatively correlated with boredom and the neutral affective state (McDaniel et al. 2007). This suggests that being in the affective states of confusion and delight may be related to learning difficult and complex issues.

The second hypothesis we tested was that the use of politeness strategies in tutorial feedback needs to adapt to the change of the social distance between the learner and pedagogical agent over time. Results show that, over time, activity of AU 12 decreases in learners who received polite feedback but increases in learners who

received direct feedback. The interaction between feedback politeness levels and AU 6 activity over time was not statistically significant. There is, however, a significant correlation between overall activity of AU 6 and AU 12. AU 12 is associated with facial expressions of joy and delight (Ekman and Friesen, 1978; McDaniel 2007). And AU 6, in addition to AU 12, is the key to the Duchenne smile, which is considered by many researchers as an indication of genuine spontaneous emotions (Ekman, Davidson and Friesen, 1990). These results suggest that the second hypothesis was only partially supported. Future analysis of student's self-report of affective states and subjective evaluation of the tutorial feedback could help clarify the influence of politeness feedback on student's affective states.

The decision to use politeness strategies is mainly based on the need to mitigate face threat and the need for efficiency. As the learner becomes more familiar with the tutor, the need to mitigate face threat decreases and the need for efficiency increases. For learners in the polite group, the use of politeness strategies may become excessive over time. For learners in the direct group, the appreciation for efficiency in the feedback may increase. This suggests that the design of politeness strategies should adapt to the change of relationship between learner and pedagogical agent. Once the social distance decreases, the lower politeness level becomes more appropriate and more efficient. One possible improvement to this study is to check how the learner's perception of social distance with the pedagogical agent changes over time.

Future work could focus on more fine-grained analysis of facial expressions, e.g. analysis of instances where AU 6 and AU 12 coincide, instead of correlating gross activities throughout the study. In the current study, we have only two data points to show how perception of politeness, through facial expressions, in tutorial feedback changes over time. Future studies that expand over weeks or months could demonstrate whether this change is linear or nonlinear, or when would be the optimal time to adjust the politeness level. As facial expression recognition and other affect recognition techniques became available and more accurate (D'Mello et al. 2007; Zeng et al. 2009), it would help informing the pedagogical agents how the feedback was received and when the politeness level needs to be updated. Future research on the politeness effect could use these technologies to dynamically adjust politeness levels and make the pedagogical agent more socially intelligent.

References

1. McLaren, B.M., DeLeeuw, K.E., & Mayer, R.E. A Politeness Effect in Learning with Web-Based Intelligent Tutors. To be presented at the 2010 American Educational Research Association (AERA) Annual Meeting, April 30 – May 4, Denver, Colorado. (2010)
2. Wang, N., Johnson, W.L. The Politeness Effect in Intelligent Foreign Language Tutoring System. In Proc. of the 9th International Conference on Intelligent Tutoring Systems (2008)
3. Wang, N., Johnson, W.L., Mayer, R.E., Rizzo, R., Shaw, E., Collins, H. The politeness effect: Pedagogical agents and learning gains. The 12th International Conference on Artificial Intelligence in Education (2005)

4. Atkinson, R. K. Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, 94, 416-427 (2002)
5. Johnson, W.L., Rickel, J., Stiles, R., and Munro, A. Integrating pedagogical agents into virtual environments. *Presence* 7(5) (1998).
6. Lester, J., Towns, S. G., Callaway, C. B., Voerman, J. L. & FitzGerald, P. J. Deictic and Emotive Communication in Animated Pedagogical Agents. In *Embodied Conversational Agents*, J. Cassell, S. Prevost, J. Sullivan, and E. Churchill (Eds.) 123-154, MIT Press, Cambridge (2000).
7. Moreno, R. Multimedia Learning with Pedagogical Agents. In R.E. Mayer (Hrsg.), *The Cambridge Handbook of multimedia learning* (S. 507 – 523). New York: Cambridge University Press (2005).
8. Baylor, A. L. The impact of pedagogical agents image on affective outcomes. In *Proc. Of the International Conference on Intelligent User Interfaces* (2005).
9. Baylor, A.L & Ryu, J. Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research*, 28 (2003)
10. Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N. AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, and J. Kay (Eds.), *Proceedings of Artificial Intelligence in Education* (pp. 47-54). Amsterdam: IOS Press. (2003)
11. Moreno, R. & Mayer, R. E. Meaningful design for meaningful learning: Applying cognitive theory to multimedia explanations. *ED-MEDIA 2000 Proceedings Charlottesville, VA: AACE Press, 747-752* (2000)
12. Moreno, R., Mayer, R. E., Spires, H., & Lester, J. The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-213 (2001)
13. McLaren, B.M., Lim, S., Yaron, D., & Koedinger, K.R. Can a Polite Intelligent Tutoring System Lead to Improved Learning Outside of the Lab? In R. Luckin, K.R. Koedinger, & J. Greer (Eds.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education*. IOS Press. 433-440 (2007)
14. Brown, P., & Levinson, S. C. *Politeness: Some universals in language use*. New York: Cambridge University Press. (1987)
15. Ekman, P. Facial expression of emotion. *American Psychologist*, 48, 384-392 (1993)
16. Ekman P. and Friesen, W. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto, CA, (1978).
17. Ekman, P. & Rosenberg, E.L., (Eds.) *What the face reveals: Basic and applied studies of spontaneous expression using the FACS*, Oxford University Press, Oxford, UK. (2005)
18. Bartlett, M., Littlewort, G. Lainscsek, C., Fasel, I., and Movellan, J. Machine learning methods for fully automatic recognition of facial expressions and facial actions. In *IEEE International Conference on Systems, Man & Cybernetics*, The Hague, Netherlands, 592-597 (2004)
19. Bartlett M.S., Littlewort G.C., Frank M.G., Lainscsek C., Fasel I., and Movellan J.R., Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia*. 1(6) p. 22-35 (2006)

20. Susskind, J.M., Littlewort, G.C., Bartlett, M.S., Movellan, J.R., and Anderson, A.K. Human and Computer Recognition of Facial Expressions of Emotion. *Neuropsychologia*. 45 (1) p. 152-162 (2007)
21. McDaniel, B. T., D'Mello, S. K., King, B. G., Chipman, P., Tapp, K., & Graesser, A. C. Facial Features for Affective State Detection in Learning Environments. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Science Society. (2007)
22. Johnson, W.L. Serious use of a serious game for language learning. In R. Luckin et al. (Eds.), *Artificial Intelligence in Education*, 67-74. Amsterdam: IOS Press. (2007)
23. Boekaerts, M. The On-Line Motivation Questionnaire: A self-report instrument to assess students' context sensitivity. In P. R. Pintrich & M. L. Maehr (Eds), *New Directions in Measures and Methods, Series in Advances in Motivation and Achievement*, 12. 77-120 (2002)
24. Littlewort, G., Bartlett, M.S. and Lee, K. Automated measurement of spontaneous facial expressions of genuine and posed pain. *Proc. International Conference on Multimodal Interfaces*, (2007)
25. Wang, N., Gratch, J. Can Virtual Human Build Rapport and Promote Learning? In *Proc. of The 14th International Conference on Artificial Intelligence in Education* (2009)
26. Ekman, P., Davidson, R.J., & Friesen, W.V. The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, 58, 342-353 (1990)
27. D'Mello, S.K., Picard, R., & Graesser, A.C. Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22(4), 53-61 (2007)
28. Zeng, Z., Pantic, M., Roisman, G.I. and Huang, T.S. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 39-58 (2009)