Real vs. acted emotional speech

Janneke Wilting, Emiel Krahmer & Marc Swerts

Communication and Cognition
Tilburg University, The Netherlands
{E.J.Krahmer/M.G.J.Swerts}@uvt.nl

10.21437/Interspeech.2006-276

Abstract

Even though the use of actors is a popular method for researching the expression of emotion, little is known about the relation between acted and real emotions. To shed some light on this, we set up a novel experiment, based on the Velten mood induction procedure, during which participants have to utter pre-defined sentences with a strong emotional content. In one group of participants, real positive or negative emotions were induced, while another group was instructed to act positive or negative while uttering Velten sentences. Results of a mood questionnaire revealed that participants in the real emotion condition, indeed felt positive or negative, depending on whether they read positive or negative sentences, while participants in the acted emotion condition felt neutral afterwards. In a second, perception experiment, it was found that acted emotions (especially negative ones) were perceived more strongly than the real emotions. This suggests that actors do not feel the acted emotion, and may engage in overacting, which casts doubt on the usefulness of actors as a way to study real emotions.
Index Terms: Emotional Speech, Audiovisual Speech, Acting.

1. Introduction

It is generally assumed that when speakers are in a particular mood or emotion this has a clear impact on how they express themselves. Most people find it relatively easy to determine a speaker's emotional state (e.g., depressed or elated) on the basis of the speaker's intonation, gestures and facial expression. There is large body of scientific research to back up this claim, especially where facial expressions are concerned (see e.g., Schmidt and Cohn 2002).

Various methods exist to study the expression and recognition of emotion, and one particularly popular method involves the use of actors which are instructed to display specific emotions. The seminal work of Darwin (1872) and Ekman (1972), for instance, is based on posed photographs of actors, and also in speech research actors are frequently used ("... and now please speak in a depressed voice"). One important advantage of using actors is that it is generally easier to instruct an actor than to elicit real emotions in participants (especially negative ones), and ethical issues are no stumbling block. But, there are disadvantages as well, mainly concerning ecological validity: it is unclear to what extent acted emotions are representative of real emotions. Given the relative importance of actors for emotion research, it is somewhat surprising that no studies (to the best of our knowledge) have systematically compared the audiovisual expression of acted and real emotions in a controlled experimental setting.

The aim of this paper is to look in more detail at the relation between audiovisual expressions of acted and real positive and negative emotions, making use of the Velten mood induction procedure (Velten 1968). A mood induction procedure (MIP) is an experimental method to elicit particular emotional states in participants. According to the meta-review of Westerman et al. (1996), the Velten method is the most widely used of such procedures, although its effectiveness differs widely across experiments. For the purpose of audiovisual speech research, the method seems highly suitable, since it involves the spoken realisation of a set of sentences with an increasingly strong emotional content.

We report on two experiments, one looking at the production of acted vs. real emotions and one looking at their perception. In the first experiment, described in section 2, we use the Velten method in a novel way to elicit both real and acted emotional speech. One group of participants is instructed to simply read the positive or negative sentences out loud (the original Velten test), while another group of participants is instructed to act as if they are in a positive or negative mood while reading sentences with an incongruous emotional content (thus, participants who acted positive read negative sentences, and participants who acted negative read positive sentences). We were interested in how people felt after doing this, which was measured using an established mood questionnaire. If the Velten method would work in our set-up, we would expect that non-acting participants who read the positive sentences feel positive afterwards, and those reading the negative sentences feel negative. The main open question is how the actors feel. There is a popular belief that displaying certain emotions leads to feeling them ("Sit all day in a moping posture, sigh and reply to everything in a dismal voice, and your melancholy lingers," James 1884), and various experimental studies have confirmed this effect (e.g., Stepper and Strack 1993). On the other hand, there is also evidence that continuously displaying a smile, e.g., in occupations requiring constant cheerfulness, definitely does not lead to a systematic positive mood (Kotchemidova 2005).

In the second experiment, described in section 3, we show film fragments of all speakers recorded in the production experiment to participants who have to rate the perceived emotion on a valency scale, ranging from very negative to very positive. Here we are primarily interested in whether or not the acted emotions are perceived in a similar way as real emotions are. If not, this would suggest that the use of actors to study real emotional speech is problematic.

2. Experiment I

2.1. Method

2.1.1. Design

The first experiment had a between participants design with condition as the independent variable (with levels: Neutral, Positive,

Act positive, Negative and Act negative) and emotional state as the dependent variable.

2.1.2. Participants

Fifty people (10 per condition) participated in the experiment, 31 female and 19 male, with an average age of 27 (range 19-52), none of the participants a (professional) actor. All were students and colleagues from Tilburg University, none involved with audiovisual speech or emotions. None of the participants objected to being filmed, and all gave written consent to use their recordings for research purposes.

2.1.3. Materials

The original stimuli set of Velten (1968) consisted of 180 sentences evenly distributed over three conditions (positive, negative and neutral). Positive and negative sentences were first literally translated in Dutch, after which they were revised to make sure they were easy to pronounce. Sentences that referred to 'specifics' (e.g., college, parents, religion) were omitted. The neutral sentences ("There is a large rose-growing center near Tyler, Texas") were replaced with comparable sentences tailored towards the Dutch situation. In the end we selected 40 sentences for each condition. We made sure that the 40 sentences in the positive and negative condition showed the same progression as the original sets of 60 sentences, from neutral ("Today is neither better nor worse than any other day") to increasingly more emotional sentences ("God I feel great!" and "I want to go to sleep and never wake up." for the positive and negative sets, respectively), to allow for a gradual build up of the intended emotional state.

2.1.4. Procedure

Participants took part one at a time. They were invited to a quiet room, where they were asked to take a seat in front of a desk on which a laptop computer was placed. The laptop was lifted 13cm from the surface so that the screen was more or less at eye level. Right above the screen a digital camera was positioned that recorded the face and upper body of the participants.

Besides the three conditions described by Velten for the induction of real emotions (POSITIVE, NEUTRAL, NEGATIVE), two acting conditions were added. In one of these, participants were shown the negative sentences and were asked to utter these as if they were in a positive mood (ACT POSITIVE), in the other, positive sentences were shown and participants were instructed to utter these in a negative way (ACT NEGATIVE).

Participants were told that the goal of the experiment was to study the effect of mood on memory recall (earlier work has revealed that the effectiveness of mood induction procedures increases when the induction serves a clear purpose, e.g., Westermann et al. 1996). The instructions, a slightly abridged version of the original instructions from Velten, were displayed on the computer screen, and participants were instructed to first silently read the texts, after which they had to read them aloud. This enabled them to practice the experimental procedure. The introduction phase was self-paced.

If the instructions were clear, the experimenter left the room and the actual experiment started. During this phase, the sentences were displayed on a computer screen for 20 seconds, and participants were instructed to read each sentence twice (once silently, then out loud). This phase lasted exactly 800 seconds (40 sentences $\times$ 20 seconds), i.e., a little over 13 minutes.

Table 1: Induced emotional state on a 7-point scale (1 = very negative, 7 = very positive) as a function of condition (standard deviations between brackets).

Condition	Emotion (s.d.)
POSITIVE	5.65 (0.63)
ACT POSITIVE	4.77 (1.23)
NEUTRAL	4.95 (0.87)
ACT NEGATIVE	4.92 (0.63)
NEGATIVE	3.85 (1.20)

Immediately following this phase, participants had to fill in a short mood questionnaire ("At this moment, I feel ...") derived from Mackie & Worth (1989) and Krahmer et al. (2004), consisting of six 7-point bipolar semantic differential scales, using the following adjective pairs (English translations of Dutch originals: happy/sad, pleasant/unpleasant, satisfied/unsatisfied, content/discontent, cheerful/sullen and in high spirits/low-spirited. The order of the adjectives was randomized; for processing negative adjectives were mapped to 1 and positive ones to 7. After filling in the questionnaire participants performed a dummy recall test, as this was supposed to be the purpose of the mood induction. The results of the recall test were not analysed. Finally, participants were debriefed and told about the real purpose of the experiment. They were given a candy bar as a token of appreciation (it has been suggested that giving a candy bar is itself a positive MIP, Westermann et al. 1996).

2.2. Results

The internal consistency of the mood questionnaire was measured using Cronbach's α and was very good ($\alpha = .92$). A t-test revealed that there was no effect of gender on the induced mood ($t(48) = .84, p = .41$), hence below we collapse data across gender. Table 1 reveals the induced emotional state on a 7-point scale (1 = very negative, 7 = very positive) as a function of condition. Clearly, the non-acting participants reading the sentences intended to induce a positive emotion feel the most positive, and non-acting participants reading the sentences intended to induce a negative emotion indeed feel most negative. Acting induces essentially the same emotion as reading the neutral sentences (i.e., a neutral one). It is interesting to observe that the standard deviations for the particapants reading the negative sentences (i.e., real negative and act positive) are twice as high as those of the other three conditions, suggesting that the effects of the negative sentence are more variable across participants than the effects of the positive sentences. A univariate analysis of variance (ANOVA) confirmed that condition had a significant effect on emotional state of the participants ($F(4, 45) = 4.65, p < .005$). A Tukey HSD post hoc analysis revealed that this effect could only be attributed to the contribution of the really positive and really negative conditions, which differed significantly from each other.

2.3. Conclusion

The first experiment revealed that the Velten method in this set-up worked very well, in the non-acting conditions the intended emotional states were indeed induced, which shows that translating the set of Velten sentences and reducing it from 60 to 40 sentences per condition did not have a negative effect on the usefulness of

Figure 1: Representative stills of real (top) and acted emotional (bottom) expressions, with on the left hand side the positive and on the right hand side the negative versions.

the method. It is highly interesting to observe that acting participants in this set-up felt neutral afterwards. In other words: the acted emotions in this experiment are not felt. The logical follow up question is how expressions of acted and real emotional speech are perceived, a question which we address in Experiment II.

3. Experiment II

3.1. Method

3.1.1. Design

The second experiment had a within participants design with condition as independent variable (with levels: Neutral, Positive, Negative, Act positive and Act negative) and perceived emotional state as the dependent variable.

3.1.2. Participants

Forty people participated (all different from those of the first experiment), 20 females and 20 males, with an average age of 36.

3.1.3. Materials

From each of the speakers in the first experiment the last sentence was selected. Depending on the condition, all speakers produce the Dutch counterparts of "God, I feel great!" (positive), "I want to go to sleep and never wake up" (negative) or "Mandarin is the official language of China" (neutral). These sentences, uttered just before filling in the mood questionnaire, arguably capture the speakers at the height of the induced emotion. The recordings were presented without sound, to avoid participants from relying on lexical cues.

3.1.4. Procedure

Participants took part one at a time. They were invited into a quite room, and asked to take place in front of a computer. Participants were told that they would see 50 speakers in different emotional states, and that their task was to rate the perceived state on a 7 point valency scale ranging from 1 (= "very negative") to 7 = ("very positive"). Participants were not informed about the fact that some of the speakers were acting. The stimuli were offered in one of two random orders, to compensate for potential learning effect. They were preceded by a number displayed on the screen indicating which stimulus would come up next, and followed by a 3 second interval during which participants could fill in their score on an answer form. Stimuli were shown only once. The experiment was preceded by a short training session consisting of three speakers (for which a different sentence was used) to make participants acquainted with the stimuli and task. If all was clear, the actual experiment started, after which there was no further interaction between participant and experimenter. The entire experiment lasted approximately 10 minutes.

Table 2: Perceived emotional state on a 7-point scale (1 = very negative, 7 = very positive) as a function of condition (standard deviations between brackets).

Condition	Perceived emotion (s.d.)
POSITIVE	4.81 (0.36)
ACT POSITIVE	4.86 (0.40)
NEUTRAL	3.52 (0.44)
ACT NEGATIVE	2.54 (0.50)
NEGATIVE	3.06 (0.40)

3.2. Results

Table 2 summarizes the results. A univariate analysis of variance (ANOVA) shows that condition has a significant effect on perceived emotional state ($F(2.66, 103.89) = 472.79, p < .001$, after a Greenhouse-Geisser correction). Post hoc analyses using the Tukey HSD method reveal that all conditions lead to a significantly different perceived mood, with the sole exception of the difference between Positive and Act positive. It is interesting to observe that the acted moods are perceived more strongly than the real ones; speakers in the Act positive condition are perceived as the most positive (although the difference with the people in a really positive mood is minimal), and speakers in the act negative condition are perceived as the most negative.

3.3. Conclusion

The perception experiment revealed that seeing speakers producing acted emotional speech leads to more extreme perceived emotion scores than seeing speakers produce real emotional speech, where the difference between acted and real emotional speech is particularly strong for the negative conditions.

4. General discussion

In this paper we described two experiments, focussing on the production and perception of real and acted emotional speech. Both studies revealed clear differences between the two: acted emotional speech is not felt, and is perceived more strongly than real emotional speech. This sheds doubt on the use of actors for emotion research, especially if the goal is to study real emotions. Contrary to what might be expected on the basis of James (1884) or Stepper and Strack (1993) but in line with, for instance, Kotchemidova (2005), displaying a certain emotion without feeling it did not have a noticeable influence on how people felt after the experiment. It might be that this is partly due to the fact that participants acting a certain emotion, had to do so with sentences expressing an incongruent emotion. In future work, we intend to redo the perception experiment with other sentences, with a less strong emotional content, to filter out possible incongruency effects

A limitation of the perception study, described in section 3, is that it only consisted of visual speech stimuli. Including auditory speech in the stimuli would have been problematic, because the lexical material would have offered very strong cues about the valency of the emotion (positive-negative) and whether it was acted or not. Interestingly, the positive acted negative sentences sound highly ironic (imagine someone saying he feels great with a depressed voice), while the negative acted positive sentences appear to be much less ironic. We are currently preparing a perception test to be carried out in the Czech republic: since Czech participants are not capable of understanding the Dutch words, they should be able to classify the perceived emotional state on the basis of both auditory and visual information, without noticing apparently ironic discrepancies between what is said and how it is said, and without taking lexical cues into account.

Finally, we would like to get a better understanding of what makes acted emotional speech different from real emotional speech. To find out, we plan to perform a second perception test; a variant of the perception test described in section 3, but now with the task for participants to guess which speakers are acting and which are not. In this way, we hope to find out which participants are good actors and which are not, and what the respective characteristic properties of acted and real emotional speech are. In addition, we want to see whether there are certain objective, acoustical and/or visual measurements which help distinguishing between acted and real emotions.

5. Acknowledgements

This paper is based on a master thesis written by the first author under supervision of the second and third author. The research was conducted as part of the VIDI-project "Functions of Audiovisual Prosody (FOAP)", sponsored by the Netherlands Organisation for Scientific Research (NWO), see `foap.uvt.nl`. Thanks to Lennard van de Laar for his help in conducting the experiments. We are very grateful to dr. Emmett Velten for supplying us with the sentences used in his original experiment.

6. References

[1] Darwin, C. (1872). *The expression of the emotions in man and animals*. 3rd edition, Oxford University Press [1998].

[2] Ekman, P. (1972). *Emotion in the human face*. Pergamon Press.

[3] James, W. (1884). What Is an Emotion? *Mind, 9*, 190–204.

[4] Kotchemidova, C. (2005). From good cheer to "Drive-By Smiling": a social history of cheerfulness. *Journal of Social History, 39(1)*.

[5] Krahmer, E., J. van Dorst and N. Ummelen (2004). Mood, persuasion and information presentation, *Information Design Journal + Document Design, 12*, 40–52.

[6] Mackie, D. and L. Worth (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57*, 27–40.

[7] Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227–256.

[8] Schmidt, K. and J. Cohn (2002). Human facial expressions as adaptations: Evolutionary questions in facial expressions, *Yearbook of Physical Anthropology, 44*, 3–24.

[9] Stepper, S. and F. Strack (1993). Proprioceptive Determinants of Emotional and Non-Emotional Feelings. *Journal of Personality and Social Psychology, 64*, 210–22.

[10] Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research & Therapy, 6*, 473–482.

[11] Westerman, R., K. Spies, G. Stahl and F. Hesse (1996). Relative effectiveness and validity of mood induction procedures, *European Journal of Social Psychology, 26*, 557–580.