The Effects of Emotionally Worded Synthesized Speech on the Ratings of Emotions and Voice Quality

Mirja Ilves[1], Veikko Surakka[1], and Toni Vanhala[1,2]

[1] Tampere Unit for Computer-Human Interaction, School of Information Sciences, FI-33014 University of Tampere, Finland
[2] VTT Technical Research Centre of Finland, P.O. Box 1300, FI-33101 Tampere, Finland
{mirja.ilves,veikko.surakka}@cs.uta.fi, toni.vanhala@vtt.fi

Abstract. The present research investigated how the verbal content of synthetic messages affects participants' emotional responses and the ratings of voice quality. 28 participants listened to emotionally worded sentences produced by a monotonous and a prosodic tone of voice while the activity of corrugator supercilii facial muscle was measured. Ratings of emotions and voice quality were also collected. The results showed that the ratings of emotions were significantly affected by the emotional contents of the sentences. The prosodic tone of voice evoked more emotion-relevant ratings of arousal than the monotonous voice. Corrugator responses did not seem to reflect emotional reactions. Interestingly, the quality of the same voice was rated higher when the content of the sentences was positive as compared to the neutral and negative sentences. Thus, the emotional content of the spoken messages can be used to regulate users' emotions and to evoke positive feelings about the voices.

Keywords: Emotions, speech synthesis, facial expression, voice quality.

1 Introduction

In human interaction, speech is the most important means to communicate ideas, intentions, and emotions, for example. For this reason, it is evident that for the development of human-computer interaction (HCI) and communication speech offers big potential. The use of synthetic speech in HCI is perhaps the most straightforward approach for creating human-like expressive verbal behavior for computers.

The importance and potential of synthesized spoken messages is highlighted in studies, which have shown that people respond to speech similarly whether it comes from a human or a computer [1]. Our approach is to support the use of speech in HCI by studying how emotional messages produced by speech synthesis affect emotional experiences and physiological responses of humans. More specifically, the focus of our research is to study the role of the verbal content in activating human emotional system.

There is evidence that the emotional content of synthetic speech has both significant effects on human emotional processing in general and potential benefits for HCI in particular. It has been found that emotionally charged messages produced by a monotonous voice can evoke emotion related subjective and physiological

S. D´Mello et al. (Eds.): ACII 2011, Part I, LNCS 6974, pp. 588–598, 2011.
© Springer-Verlag Berlin Heidelberg 2011

responses in listeners [2,3]. Further, cognitive processing can be facilitated by giving emotionally worded feedback during a computerized problem-solving task [4].

On the other hand, we have suggested that the naturalness of voice affects the experienced emotions. Synthesized speech can be created with different techniques that sound more or less humanlike. We have found that when emotional messages were delivered by two synthesizers, only the more humanlike synthesizer evoked emotion related facial electromyographic (EMG) responses as well as more emotion-related variation in the ratings of arousal and in pupil responses [2,3]. In addition to synthesis technique, the naturalness of prosody is a central component for natural sounding voice. The naturalness of voice can easily be increased by adding prosodic variation to speech that would otherwise be monotonous. This enables us to study whether emotional effects could be enhanced by increasing the prosodic variation of the speech. In the present research, emotionally worded sentences were produced by unit selection technique with a monotonous and a prosodic speaking style.

Naturalness, clearness and pleasantness are central factors when people rate the quality of speech synthesis [5]. Adding prosodic variation to the synthesized speech probably enhances the quality of voice, but it is interesting to study if the ratings of voice quality are affected also by the content of the message. Fogg and Nass [6] found that interaction with a computer that flatters was rated more positively than the interaction with a computer that gave generic feedback. Thus, the second interest of the present study is related to the question: Is the verbal content of the spoken message so powerful that it affects the ratings of voice quality, even though the voice remains similar? If a synthesizer speaks in a positive way about a listener, does the listener rate the voice more positively than when the synthesizer produces negative statements about the listener? Does this happen even when there is no real interaction between the human and the computer?

Dimensional theory of emotions defines emotions through a set of dimensions [7]. The most important and widely used dimensions are valence, which refers to the pleasantness of the experienced emotion and arousal, which relates to the experienced level of activation.

Emotions cause changes both in subjective experiences and physiological activity. For example, there is evidence that the activation of facial muscles is related to emotional valence. It has been shown that there is a linear effect of valence on activity over corrugator supercilii [8]. In other words, the activity of corrugator supercilii increases during negative emotions and relaxes during positively valenced emotions. Based on this, we chose to investigate emotion related physiological responses by measuring the activity of corrugator supercilii in the present study. However, it is important to remember when viewing the results that corrugator activity may be influenced also, for example, by cognitive factors like concentration [9].

In summary, the aim of the present research was to study how synthetic verbal stimulation with emotional content might affect the perceiver. First aim was to study, if there is a difference between the emotional effects of same messages delivered either without or with prosodic variation. Second aim was to study whether the ratings of voice quality are affected by the emotional content of the spoken message. First, participants listened to a set of emotionally worded sentences produced by a monotonous and a prosodic tone of voice, while their corrugator EMG activity was measured. Following this, ratings of emotional experiences and speech quality were collected.

2 Methods

2.1 Participants

31 volunteers from an introductory course of Computer Science participated in the study. Data from three subjects was discarded, because post-experiment interview revealed that they were aware of the facial muscle activity being measured. Thus, data from 28 subjects were used for the analysis (21 females, mean age 22.93 years, range 18-41 years). The participants were native speakers of Finnish, and had normal hearing by their own report. The participants were compensated with a movie ticket for their participation.

2.2 Equipment and Physiological Measurement

Stimuli were created with Finnish-speaking speech synthesizer Bitlips' Unit Selection [10]. The stimulus presentation and rating were controlled by E-Prime© [11] stimulation software running on a PC computer with a Windows XP operating system. The stimuli were presented via earphones.

Facial EMG was registered from the left side of the face above the corrugator supercilii muscle site using Ag–AgCl sintered electrodes. The ground electrode was over the mastoid bone and an active reference on the forehead close to the center of the hairline. Guidelines of Fridlund and Cacioppo [12] were followed. EMG was measured using a NeXus-10 physiological monitoring device (Mind Media B.V.) that was connected to a laptop computer using a wireless Bluetooth communications link. The sampling rate was 2048 Hz. Analog high-pass filter of .5 Hz was used and EMG was further digitally pass-band filtered (4-th order Butterworth) from 20 to 500 Hz.

2.3 Stimuli

Vocal stimuli consisted of 12 different sentences[1] with emotional content. Emotional categories were happiness, anger, and neutral. Happiness and anger were selected because they are located on the opposed ends of valence scale. Emotional words expressing happiness and anger were selected and translated to Finnish from the list of causative emotions by Johnson-Laird & Oatley [13]. In addition, neutral sentences were created so that the sentences in different categories matched for word length.

The sentences were produced by a male voice of synthesizer with both a monotonous and a prosodic tone of voice. The monotonous tone of voice means that fundamental frequency (F0) variation was set as flat as possible. The prosodic tone of voice means that variation in F0 was not manipulated, that is, variation in F0 was not flattened nor any emotional cues were added. Besides the variation of F0, the voice parameters of the sentences were similar to each other as shown in Table 1.

---

[1] Sentences (roughly translated from Finnish): Happiness: I am content with you, You please me, You make me delighted, You make me exhilarated; Anger: I am irked by you, You irritate me, You make me annoyed, You make me infuriated; Neutral: You seem to be normal, You are average, You seem to be decent, You seem to be ordinary.

Table 1. ANOVA results for the duration (in seconds), loudness (dB), fundamental frequency (F0), and variability of fundamental frequency (F0 sd) of the stimulus sentences produced by the monotonous and prosodic tone of voice

	Duration (s)	dB	F0 (Hz)	F0 (Hz) sd
Monotonous	1,33	72,63	89,37	3,31
Prosodic	1,33	73,21	89,39	7,94
ANOVA	$F(1, 18) = 0.00$, ns	$F(1, 18) = 1.26$, ns	$F(1, 18) = 0.00$, ns	$F(1, 18) = 152.52$, $p< .001$

2.4 Experimental Procedure

First, the sound attenuated and electromagnetically shielded laboratory was introduced and the participant was seated in a chair. She/he was explained that the aim was to study how people react to synthesized speech. The participant was told a cover story that the purpose of the electrodes was to measure the humidity of her/his skin, and then the electrodes were attached.

To familiarize the participants to listening to synthetic speech, three speech samples spoken by three different speech synthesizers were presented before the actual experiment. Then, the participant was instructed to relax, listen carefully with an open mind the sentences, and to fixate to a small cross in the center of the screen whenever the cross was within sight. Finally, the lights of the laboratory were dimmed and the experimental phase began.

The stimulus sentences were presented randomly with fifteen seconds pause between the stimulus presentations. Fixation-cross appeared five seconds before the stimulus onset and disappeared five seconds after the stimulus offset. The presentation of the stimulus sentences lasted approximately 8 minutes.

After the experimental phase, the electrodes were detached and the participant gave the ratings of the stimuli. First, the participant rated her/his emotional experiences evoked by the stimuli on nine-point bipolar scales. The valence-scale ranged from negative to positive and the arousal-scale from calm to aroused. Secondly, the participant rated on five-point scales how pleasant, natural, and clear the voice was. These questions were selected from a modified mean opinion score (MOS) scale [5]

Before the rating sessions, the scales were explained to the participant. The scales were presented in the computer screen and a keyboard was used to give the ratings. After the ratings, the participant was interviewed whether she/he was aware that her/his facial muscle activity had been measured. Finally, the participant was debriefed about the purpose of the study, and she/he gave a written consent.

2.5 Data Preprocessing and Analysis

EMG responses were extracted by averaging rectified sample values. A 500 ms pre-stimulus baseline correction was performed. Mean EMG responses of three seconds from the stimulus offset were analyzed.

Subjective data was analyzed using repeated measures ANOVAs. Greenhouse-Geisser adjusted degrees of freedom were used when necessary. For the multiple post-hoc comparisons, Bonferroni corrected p-values were used.

3 Results

3.1 Ratings of Emotional Experiences

Figure 1 shows the mean ratings and standard error of the means (SEM) for valence and arousal. 2 x 3 two-way (voice x emotion category) ANOVAs were performed separately for the ratings of valence and arousal. For the ratings of valence, there was a statistically significant main effect of emotion category, $F(1, 36)= 42$, $p< .001$. The main effect of voice and the interaction of voice and category were not statistically significant. Post-hoc pairwise comparisons showed that the ratings were significantly lower for the negative sentences than for the neutral, $t(27)= -3.98$, $p< .001$ and positive sentences, $t(27)= -6.98$, $p< .001$. The ratings for the positive sentences were significantly higher than for the neutral sentences, $t(27)= 6.59$, $p< .001$.

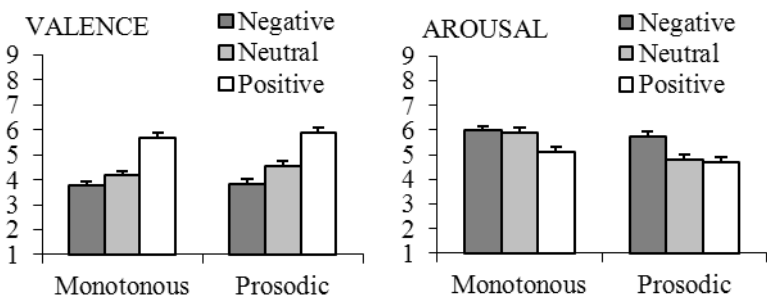

Fig. 1. Mean ratings (and SEM) of valence and arousal

For the ratings of arousal, ANOVA revealed a statistically significant main effect of voice, $F(1, 27)= 12.79$, $p< .01$ and emotion category, $F(2, 54)= 10.47$, $p< .001$. Because the interaction effect was also significant, $F(2, 54)= 6.76$, $p< .01$, one-way ANOVAs with emotion category as a factor were conducted separately for both voices. The results are reported in Table 2.

Table 2. Results of the effect of the voice on the ratings of arousal

Tone of voice	ANOVA	Pairwise comparisons
Monotonous	$F(2, 54) = 7.89$, $p< .001$	Negative > Positive $t(27) = 3.54$, $p< .01$ Neutral > Positive $t(27) = 2.98$, $p< .05$
Prosodic	$F(2, 54) = 11.35$, $p< .001$	Negative > Positive $t(27) = 4.30$, $p< .001$ Negative > Neutral $t(27) = 4.03$, $p< .01$

In addition, pairwise comparisons between the voices were conducted separately for the different emotion categories. The neutral sentences produced by the monotonous voice were rated as more arousing than the neutral sentences spoken by the prosodic tone of voice, $t(27)= 4.61$, $p< .001$. The arousal ratings of the positive sentences produced by the monotonous voice were higher than the arousal ratings of the positive sentences produced by the prosodic tone of voice, t(27)= 2.70, p< .05. The difference between the arousal ratings of the negative sentences was not statistically significant.

3.2 Ratings of Voice Quality

The average ratings of the voice are shown in Figure 2. Three 2 x 3 two-way (voice x emotion category) ANOVAs were performed separately for each ratings. For the ratings of pleasantness, there was a statistically significant main effect of voice, F(1, 27)= 57.84, p< .001 and emotion category, F(2, 54)= 25.89, p< .001, and a significant interaction of voice and category, F(2, 54)= 11.40, p< .001. For the ratings of naturalness, there was a statistically significant main effect of voice, F(1, 27)= 76.29, p< .001 and emotion category, F(2, 54)= 25.54, p< .001, and a significant interaction of voice and category, F(2, 54)= 7.89, p< .001. Also, for the ratings of clearness, there was a statistically significant main effect of voice, F(1, 27)= 62.38, p< .001 and emotion category, F(2, 54)= 39.57, p< .001, and a significant interaction of voice and category, F(2, 54)= 6.82, p< .01.

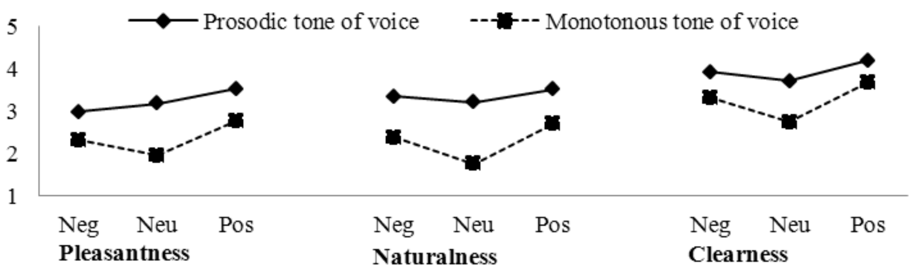

Fig. 2. Mean ratings of pleasantness, naturalness, and clearness of the voice separately for both voices and each emotion category

Table 3. Results of the effect of the emotion category on the ratings of the pleasantness, naturalness, and clearness of the monotonous voice

Rating scale	ANOVA	Pairwise comparisons
Pleasantness	F(2, 54) = 29.06, p < 0.001	Positive>Neutral t(27) = 7.22, p < .001 Positive>Negative t(27) = 4.50, p < .001 Negative>Neutral t(27) = 3.36, p < .01
Naturalness	F(2, 44) = 40.78, p < 0.001	Positive>Neutral t(27) = 12.05, p < .001 Positive>Negative t(27) = 2.78, p < .05 Negative>Neutral t(27) = 5.14, p < .001
Clearness	F(2, 54) = 48.33, p < 0.001	Positive>Neutral t(27) = 9.96, p < .001 Positive>Negative t(27) = 3.95, p < .01 Negative>Neutral t(27) = 5.60, p < .001

Pairwise comparisons between the voices showed that the prosodic tone of voice was rated as more pleasant, t(28)= 7.61, p< .001, natural, t(28)= 8.73, p< .001, and clear, t(28)= 7.90, p< .001 than the monotonous voice.

Because of the significant interaction, one-way ANOVAs with emotion category as a factor were performed separately for the both voices. The emotional content of the sentences affected the ratings of voice quality as can be seen from Tables 3 and 4.

Table 4. Results of the effect of the emotion category on the ratings of the pleasantness, naturalness, and clearness of the prosodic tone of voice

Rating scale	ANOVA	Pairwise comparisons
Pleasantness	F(2, 54) = 13.10, p < .001	Positive>Neutral t(27) = 3.28, p < .01 Positive>Negative t(27) = 4.62, p < .001
Naturalness	F(2, 54) = 2.38, p > .05	
Clearness	F(2, 54) = 9.24, p < .001	Positive>Neutral t(27) = 4.13, p < .001 Positive>Negative t(27) = 3.05, p < .05

3.3 Electromyographic Responses

Mean corrugator responses to stimulus categories are presented in Figure 3. Interestingly, the mean corrugator response to all stimulus categories was relaxing. First, the effect of the voice on corrugator muscle activity was tested. Pairwise comparison between the voices did not reveal a statistically significant difference between the voices, t(27) = 1.67, ns. Because the degree of the relaxation seemed to be divergent to different stimulus categories, the activity change of the corrugator muscle was compared to the baseline level. Statistically significant relaxations from the pre-stimulus baseline were found by one-sample t-test for the neutral sentences, t(27) = -2.15, p < .05. Relaxation from the pre-stimulus baseline for the positive sentences approached statistical significance, t(27) = -1.894, p < .07. Mean corrugator response to negative sentences did not differ significantly from the baseline level, t(27) = -0.93, p > ns.

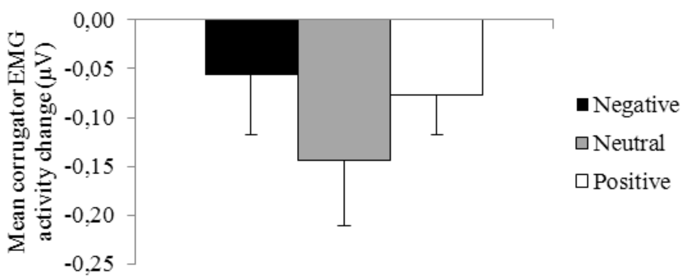

Fig. 3. Mean corrugator supercilii EMG responses (and SEM) to negative, neutral, and positive sentences during three-second period after stimulus offset

4 Discussion

The ratings of valence were significantly affected by the emotional contents of the sentences. The ratings were significantly different for all stimulus categories. The tone of voice had no significant effect on the ratings of valence. In line with the valence ratings, corrugator supercilii activity was not affected by the tone of voice. Instead, frowning activity significantly decreased from the pre-stimulus baseline level during the three seconds following a neutral stimulus. After negative and positive sentences, frowning activity was not significantly different from the baseline.

The ratings of arousal were affected by the voice. The neutral and positive sentences produced by the monotonous voice evoked higher ratings of arousal than the positive and neutral sentences produced by the prosodic tone of voice.

The tone of voice had an effect also on the ratings of the quality of the voice. The prosodic tone of voice was rated as more pleasant, natural, and clear than the monotonous voice. In addition, the verbal content of the sentences had interestingly an effect on the ratings of voice quality. The prosodic tone of voice was rated as more pleasant and clear when the content of the sentence was positive compared with the neutral and negative sentences. The monotonous tone of voice was rated similarly as more pleasant and clear, and also more natural, when the content of the sentence was positive in comparison to the neutral and negative sentences. The ratings of monotonous voice were also more positive when the content of the sentence was negative in comparison to the sentences with neutral content.

The current findings support the previous findings that emotional responses can be evoked by the lexical emotional content of the spoken messages [2,3]. At the present study, the ratings of valence were significantly different for all emotion categories but the ratings were not affected by the tone of voice. We have previously found [14] that even short emotional words produced by a very machinelike voice can evoke congruent emotion-related ratings in the participants, but the experienced valence and approachability can be enhanced when the words are spoken by a more humanlike voice. On the other hand, at the previous study, the differences between two concatenative synthesizers (i.e. more humanlike voices) were not significant. Thus, it seems that, when the voice is experienced humanlike enough, experienced valence is not enhanced even if the naturalness of the voice is increased by using the technique that sounds even more humanlike or increasing the prosodic variation of the voice. There are similar findings also for human speech. In the study of Bertels et al. [15], people rated the valence of positive words similarly when the words were uttered by a neutral tone of human voice compared with the emotionally congruent tone of human voice.

However, the arousal of the sentences was rated significantly differently depending on the tone of voice. The arousal of the neutral and positive sentences was rated significantly higher when the sentences were produced by the monotonous voice compared with the prosodic tone of voice. When we look at the averages of the ratings, it can be seen that the ratings of the positive and neutral sentences produced by the prosodic tone of voice were actually more emotion-relevant than the positive and neutral sentences produced by the monotonous voice. That is because the arousal of the neutral sentences produced by the prosodic tone of voice was judged closer to neutral experience (the center of the scale) than the arousal of the neutral sentences produced by the monotonous tone of voice. Further, the arousal ratings of the positive sentences produced by the monotonous voice were very close to neutral experience whereas the arousal ratings of the positive sentences produced by the prosodic tone of voice were more different from the neutral experience and were judged a little bit calming.

It has been shown that perceiving emotionally negative stimulus increases corrugator supercilii EMG activity and emotionally positive stimulus decreases the activity [8]. At the present study, the average corrugator response during three seconds period after stimulus offset was below pre-stimulus baseline level for all emotion categories. Relaxing corrugator response to speech synthesis has also been found in previous studies [2,4]. Significant decrease from the pre-stimulus baseline was found for the neutral sentences, whereas the activity of the muscle site did not differ significantly from the pre-stimulus baseline after the negative and positive sentences. Thus, the current EMG responses do not seem to reflect emotional reactions. Even though the correlation between the facial muscle activity and emotion has been frequently found, there exist also studies where this correlation has not been found [see 16]. It has also been reported that the activity of corrugator is associated with cognitive processing, like mental effort [17] and concentration [9]. It might be that corrugator activity at the present research did not reflect emotions but cognitive processing instead. Because the corrugator activity decreased significantly after the neutral sentences but not after the negative and positive sentences, it may be that the participants concentrated less on the processing of the neutral sentences than the processing of the emotionally loaded sentences. This is supported by memory studies, where it has been found that emotionally valenced words are remembered better than nonvalenced words, and this has suggested resulting, for example, from the facilitated attention to emotional words [18].

One of the most interesting findings of the current study was that the content of the spoken message affected how people rated the voice quality. In spite of the fact, that the same voice produced negative, neutral, and positive sentences and the participants were instructed to specifically rate the voice, the sentences with positive content evoked more positive ratings of the voice quality than the negative or neutral sentences. This was true for both prosodic styles. In addition, when the voice was monotonous, even the negative emotional content of the sentence had pronounced effects on the ratings of voice quality as compared to the neutral content. It has been previously found that the interaction with computer is rated more positively when the computer sincerely or insincerely praises the participant than when the computer gives only some generic feedback [6]. Our study showed that the positive content has so strong effect on people that just listening to the positive comments, without any interaction, affects the ratings of voice.

In summary, our results confirmed previous findings that emotional responses can be evoked by the lexical content of the synthetic speech. Speech itself seems to be so powerful means of communication that even the monotonous tone of voice can be used quite effectively to communicate the emotional messages. In addition to previous findings that the emotional effects of emotionally worded messages can be enhanced through the synthesized nature of the speech, emotional effects can further be enhanced by increasing the naturalness of the prosodic variation. However, because the voice itself was quite natural, increasing the naturalness through the prosodic variation did not seem to have great effects on the emotional responses. Instead, the pleasantness, naturalness, and clearness of the voice were rated clearly higher when the prosodic variation was more natural compared with the monotonous speaking style. Further, the impression of the voice quality was affected by the content of the spoken message. Thus, the emotional content of the spoken messages in general can be used to regulate users' emotions and positive language in particular can be used to enhance positive feelings about the voices.

Acknowledgments. The authors would like to thank all the participants of the study. This research was supported by the Graduate School in User-Centered Information Technology (UCIT).

References

1. Nass, C., Brave, S.: Wired for speech: How voice activates and advances the human-computer relationship. The MIT Press, Cambridge (2005)
2. Ilves, M., Surakka, V.: Subjective and physiological responses to emotional content of synthesized speech. In: Magnenat-Thalmann, N., Joslin, C., Kim, H. (eds.) Proc. of CASA, pp. 19–26 (2004)
3. Ilves, M., Surakka, V.: Emotions, Anthropomorphism of Speech Synthesis, and Psychophysiology. In: Izdebski, K. (ed.) Emotions in the Human Voice. Culture and Perception, vol. III, pp. 137–152. Plural Publishing Inc., San Diego (2009)
4. Partala, T., Surakka, V.: The effects of affective interventions in human-computer interaction. Interacting with Computers 16(2), 295–309 (2004)
5. Viswanathan, M., Viswanathan, M.: Measuring speech quality for text-to-speech systems: development and assessment of a modified mean opinion score (MOS) scale. Computer Speech and Language 19(1), 55–83 (2005)
6. Fogg, B.J., Nass, C.: Silicon sycophants: the effects of computers that flatter. International Journal of Human-Computer Studies 46(5), 551–561 (1997)
7. Bradley, M.M., Lang, P.J.: Measuring emotions: the self-assessment manikin and the semantic differential. Journal of Behavioral Therapy and Experimental Psychiatry 25(1), 49–59 (1994)
8. Larsen, J.T., Norris, C.J., Cacioppo, J.T.: Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. Psychophysiology 40(5), 776–785 (2003)
9. Cacioppo, J.T., Petty, R.E., Morris, K.J.: Semantic, evaluative, and self-referent processing: memory, cognitive effort, and somatovisceral activity. Psychophysiology 22(4), 371–384 (1985)
10. Bitlips, `http://www.bitlips.fi/index.en.html` (retrieved March 30, 2011)
11. Schneider, W., Eschman, A., Zuccolotto, A.: E-Prime User's Guide. Psychology software Tools Inc., Pittsburgh (2002)
12. Fridlund, A.J., Cacioppo, J.T.: Guidelines for human electromyographic research. Psychophysiology 23(5), 567–589 (1986)
13. Johnson-Laird, P.N., Oatley, K.: The language of emotions: an analysis of a semantic field. Cognition and Emotion 3(2), 81–123 (1989)
14. Ilves, M., Surakka, V.: Subjective emotional responses to synthesized speech with lexical emotional content: The effect of the naturalness of the synthetic voice. Submitted for publication in Behaviour & Information Technology (under review)

15. Bertels, J., Kolinsky, R., Morais, J.: Norms of emotional valence, arousal, threat value and shock value for 80 spoken french words: comparison between neutral and emotional tones of voice. Psychologica Belgica 49(1), 19–40 (2009)
16. Tassinary, L.G., Cacioppo, J.T.: The skeletomotor system: Surface electromyography. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (eds.) Handbook of Psychophysiology, 2nd edn., pp. 163–199. Cambridge University Press, New York (2000)
17. Van Boxtel, A., Jessurun, M.: Amplitude and bilateral coherency of facial and jaw-elevator EMG activity as an index of effort during a two-choice serial reaction task. Psychophysiology 30(6), 589–604 (1993)
18. Doerksen, S., Shimamura, A.P.: Source memory enhancement for emotional words. Emotion 1(1), 5–11 (2001)