Available online at www.sciencedirect.com
ScienceDirect
Cognition 104 (2007) 1–18
COGNITION
www.elsevier.com/locate/COGNIT

Spontaneous facial mimicry in response to dynamic facial expressions ☆

Wataru Sato [a,*], Sakiko Yoshikawa [b]

[a] *Department of Psychology, Graduate School of Letters, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan*

[b] *Department of Cognitive Psychology in Education, Graduate School of Education, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan*

Received 4 January 2006; revised 3 May 2006; accepted 5 May 2006

Abstract

Based on previous neuroscientific evidence indicating activation of the mirror neuron system in response to dynamic facial actions, we hypothesized that facial mimicry would occur while subjects viewed dynamic facial expressions. To test this hypothesis, dynamic/static facial expressions of anger/happiness were presented using computer-morphing (Experiment 1) and videos (Experiment 2). The subjects' facial actions were unobtrusively videotaped and blindly coded using Facial Action Coding System [FACS; Ekman, P., & Friesen, W. V. (1978). *Facial action coding system.* Palo Alto, CA: Consulting Psychologist]. In the dynamic presentations common to both experiments, brow lowering, a prototypical action in angry expressions, occurred more frequently in response to angry expressions than to happy expressions. The pulling of lip corners, a prototypical action in happy expressions, occurred more frequently in response to happy expressions than to angry expressions in dynamic presentations. Additionally, the mean latency of these actions was less than 900 ms after the onset of dynamic changes in facial expression. Naive raters recognized the subjects' facial reactions as emotional expressions, with the valence corresponding to the dynamic facial expressions that the subjects were viewing. These results indicate that dynamic facial expressions elicit spontaneous and

☆ This manuscript was accepted under the editorship of Jacques Mehler.

* Corresponding author. Tel./fax: +81 75 771 2896.
E-mail address: L50158@sakura.kudpc.kyoto-u.ac.jp (W. Sato).

doi:10.1016/j.cognition.2006.05.001

rapid facial mimicry, which functions both as a form of intra-individual processing and as inter-individual communication.

Keywords: Facial mimicry; Dynamic facial expressions; Mirror neuron

1. Introduction

Communication through facial expressions of emotion plays an important role in social coordination (Keltner & Kring, 1998). Through the evolutionary process, facial expressions would have helped humans take collective actions during danger and forming intimate relationships with other individuals. Consistent with this idea, psychophysiological studies using facial electromyography (EMG) indicate that facial expressions elicit facial muscular activity congruent with the presented facial expressions. For example, Dimberg (1982) showed that mere photographic presentations of angry and happy facial expressions induced spontaneous corrugator supercilii muscle activity (brow lowering actions, prototypical in angry facial expressions) and zygomatic major muscle activity (lip corner pulling actions, prototypical in happy facial expressions), respectively. This facial muscular activity may be interpretable as mimicking behavior or "facial mimicry" (Hess, Philippot, & Blairy, 1999). Dimberg, Thunberg, and Elmehed (2000) reported that facial EMG activity occurred even without awareness of the specific facial expression, confirming the spontaneous nature of the responses. This facial reaction occurs rapidly; Dimberg and Thunberg (1998) showed that facial EMG activity occurred after only 500 ms of exposure to the facial pictures. These data imply that facial muscle activity that may relate to facial mimicry occurs spontaneously and rapidly in response to facial expressions.

However, there is little evidence as to whether the facial muscle activity revealed by EMG recordings is externally visible as facial mimicry. This point is crucial because if overt facial mimicking occurs when perceiving facial expressions of emotion, then this facial motor activity could function not only in intra-individual processing, such as empathic understanding, but also in inter-individual communication. Although developmental studies have demonstrated that neonates exhibit overt facial mimicry of adult facial expressions (Meltzoff & Moore, 1977; Field, Woodson, Greenberg, & Cohen, 1982), the visibility of facial activity has not been explicitly measured in EMG studies with adult subjects. As these facial EMG amplitude changes are very subtle (a few microvolts), facial muscle activities may not be visible (cf. Cacioppo, Petty, Losch, & Kim, 1986).

With regard to the mechanism of facial mimicry, recent neuroscientific evidence provides the clue, pointing to the involvement of motor-related brain areas in social communication. Single-unit recording studies in monkeys have revealed that specific neurons in the ventral premotor cortex (area F5) discharge both when the monkey performs specific hand actions and when it observes experimenters performing similar actions; These neurons have been named "mirror neurons" (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). A recent study with monkeys revealed that the neurons in this region discharge during the observation of communicative facial actions (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003). Neuroimaging studies have confirmed the existence of the mirror neuron system in humans. Buccino and his colleagues (Buccino et al., 2001, 2004) showed that observing videotaped mouth actions, as compared to the observation of static faces, activated the ventral premotor area, centered to the pars opercularis of the inferior frontal gyrus of Brodmann area (BA) 44, an area that has been proposed as the human homologue of area F5 (Rizzolatti & Arbib, 1998).

A recent neuroimaging study (Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004) extended the notion of the mirror neuron system in humans into the processing of facial expressions of emotion. In this study, brain activity in response to dynamic facial expressions of fear and happiness was compared to activity in response to static expressions and dynamic mosaic images. Results revealed that, in addition to areas such as the temporal cortex, the pars opercularis of the inferior frontal gyrus (BA 44), was more active in response to the dynamic facial expressions than it was to both control stimuli that were common to both expressions. This suggests that the mirror neuron system in the ventral premotor cortex may be automatically activated when processing facial expressions of emotion, especially if the expressions are dynamic.

This neuroscientific evidence indicates that the mirror neuron system is specifically involved in the processing of dynamic actions relative to static actions. Such evidence further suggests the possibility that the subtlety of facial reactions indicated in previous facial EMG literature may have been a consequence of using static facial expressions as stimuli in almost all of the studies undertaken to date, with the exception of a few studies such as that conducted by McHugo, Lanzetta, Sullivan, Masters, and Englis (1985). We hypothesized that facial mimicry may evidently occur, in the externally visible form, in response to the dynamic facial expressions of emotion.

In the present study, we conducted two experiments to investigate whether spontaneous, externally visible facial mimicry occurs by videotaping subjects' facial reactions while they were passively observing dynamic or static facial expressions. A video prompter system, an apparatus generally used in television studios, was used to videotape the subjects' facial reactions. This allowed us to videotape subjects' facial expressions while they were unaware of the presence of a video camera. To analyze the facial reactions, scorers blindly scored the facial movements using the Facial Action Coding System (FACS; Ekman & Friesen, 1978). Action unit (AU) 4 (brow lowering, prototypical facial actions in angry expressions) and AU 12 (lip corner pulling, prototypical facial actions in happy expressions) were evaluated. To present the dynamic facial expressions, we used the computer-morphing techniques in Experiment 1 and videos of natural facial expressions in Experiment 2 because these methods both have advantage and disadvantages and are complementary to each other (Sato & Yoshikawa, 2004). We prepared facial expressions of anger and happiness to represent positive and negative emotional valence. We used the apex images of the dynamic facial expressions under static conditions. We predicted that specific facial action patterns, interpretable as facial mimicry, would occur spontaneously in dynamic presentations, but not necessarily in static presentations.

2. Experiment 1

In Experiment 1, we used a computer-morphing technique to create dynamic emotional facial expressions (Fig. 1), as in a previous neuroimaging study (Sato et al., 2004). Stimuli consisted of facial images transforming seamlessly from neutral to emotional expressions. An advantage of this method is that it allows us to strictly control the spatial and temporal parameters of facial expressions, relative to other methods such as videos of real facial expressions; however, a disadvantage of the method is that changes are artificial. In addition to analyses of the occurrence of subjects' facial reactions for each trial, temporal control of the stimuli allowed us to analyze the facial reactions frame by frame for latencies. We predicted that the specific and rapid facial reactions could occur in dynamic presentations.

2.1. Methods

2.1.1. Subjects

Eighteen volunteers (9 females and 9 males; mean age, 21.9 years) participated in this experiment. All the subjects were right-handed, and had normal or corrected-to-normal visual acuity. Although two additional subjects actually participated, their data were not analyzed due to large numbers of artifacts (see Section 2.1.6).

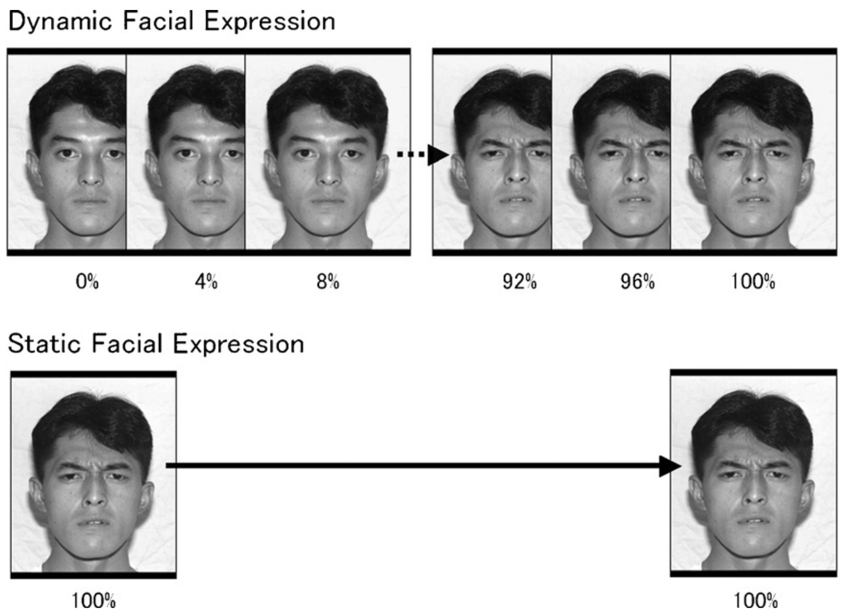

Fig. 1. Stimulus presentations in the dynamic facial expression condition (upper) and static facial expression condition (lower) in Experiment 1.

2.1.2. Experimental design

The experiment was constructed as a within-subjects two-factorial design, with presentation condition (dynamic/static) and expression (angry/happy) as the factors.

2.1.3. Stimuli

Raw materials were color photographs of the angry, happy, and neutral expressions of two females and two males. These stimuli were chosen from a database of facial expressions containing images of facial expressions posed by more than 50 Japanese models. Preliminary ratings from subjects who did not take part in this experiment confirmed that the stimuli clearly displayed the target emotions relative to other basic emotions. None of the faces were familiar to any of the subjects.

Computer animation clips from these photos served as dynamic expression stimuli. First, 24 intermediate images in 4% steps between the neutral (0%) and emotional (100%) expressions were created using computer-morphing techniques (Mukaida et al., 2000) implemented on a Linux computer. Fig. 1 displays an example of the stimulus sequence. Next, a total of 26 images (i.e., one neutral image, 24 intermediate images, and the final emotional image) were presented in succession to create a moving clip. Each image was presented for 40 ms, and the first and last images were presented for 280 ms; thus, each animation clip was presented for 1520 ms. This change in speed has been shown to sufficiently represent natural changes in the dynamic facial expressions of anger and happiness (Sato & Yoshikawa, 2004).

The emotional expressions that corresponded to the final images (100%) in the dynamic condition were prepared for the static expression stimuli and presented for 1520 ms.

2.1.4. Apparatus

Experimental events were controlled by a program written in Visual C++5.0 and implemented on a computer (Inspiron 8000, Dell) with a Microsoft Windows operating system.

We used a video prompter system (CWP10H, Canon) to present stimuli and to unobtrusively videotape the subjects' facial reactions. The system consisted of a 15-in. TFT monitor, a two-way mirror, and other interfaces. The monitor was tilted up (the subjects could not see it because of a curtain), and the screen was reflected onto a two-way mirror mounted at a 45° angle to the screen. The mirror reflected the screen, creating a viewing situation similar to that of an ordinary 15-in. monitor. The distance between the mirror and the subjects' eyes was about 0.6 m. The monitor screen was set for a 600 vertical × 800 horizontal pixels resolution, 16-bit color, and a frame refresh rate of 75 Hz. The stimuli were presented at 250 vertical × 200 horizontal pixels, which subtended a visual angle of about 9.2° vertically × 7.3° horizontally. The background was held constant with a gray color.

A video camera (DSR-PD150, SONY) was placed behind the two-way mirror. The video camera recorded the subjects' full faces at the rate of 30 frames per second (i.e., 33.3 Hz).

2.1.5. Procedure

Experiments were conducted individually in a chamber room. Upon arrival, subjects were told that the experiment concerned the evaluation of faces, and that they would be watching some facial stimuli and later answering some questions about them. The subjects relaxed for 10 min to allow general adaptation to the experimental room, after which time they viewed instructions displayed on the monitor that explained the stimulus presentation.

The subjects participated in a total of 32 trials, consisting of two presentations of each stimulus (with eight trials each of dynamic angry, dynamic happy, static angry, and static happy expressions). The order of stimulus presentation was randomized. Subjects were given a few practice trials to familiarize themselves with the stimulus presentation before data collection. To avoid habituation and drowsiness, subjects had a short rest period halfway through the trials. Throughout the stimulus presentations, the subject's full face was continuously videotaped by the video camera behind the screen, without the subject being aware of the videotaping.

In each trial, a warning tone was first presented for 100 ms. Simultaneous with the onset of the tone, a fixation point (the picture with a small "+" in a gray color on a white background and of the same size as the stimulus) was presented at the center of the screen for 520 ms. Then, the stimulus was presented for 1520 ms. After stimulus presentation, the screen was filled with a gray color as an inter-trial interval, which varied randomly from 3000 to 7000 ms. The onset of the warning tone was considered in the analyses to be the trigger signaling the beginning of each trial.

After viewing all of the images, the subjects were asked to respond to open-ended questions about the stimuli. These data are not reported here because the performance of this task was irrelevant to the purposes of the study.

Finally, the subjects were interviewed to determine whether they had been aware that their faces had been videotaped. This process confirmed that all of the subjects had been unaware of the videotaping. Debriefing was conducted several weeks later to prevent subjects from communicating the purpose of the experiment to other subjects. Subject permission to use the videotapes for analysis was requested and granted in all cases. For subjects whose data are reported here as examples, additional permission was requested later, and written consent was obtained.

2.1.6. Data analysis

2.1.6.1. Preprocessing. To analyze the subjects' facial reactions, the videotapes were digitized at 30 frames per second into a computer (Dimension 8100, Dell). Then, facial reaction in response to each stimulus was sampled for 3000 ms; the response data consisted of pre-stimulus baseline data for 500 ms (the fixation point was presented) and the data for 2500 ms after stimulus onset. The first trials for all subjects were not included in the analysis because the subjects' reactions could have incorporated orienting or startle responses.

Because subjects were not told that their faces were being videotaped, the data included various types of artifacts: covering a part of his/her face with his/her hand, resting his/her chin on his/her hand (this could prevent the movement of the lower face), swinging his/her head, yawning, eye closing, mumbling, and displaying facial expressions in the pre-stimulus period. Based on examination of the above, we defined criteria for artifacts, which were then used by a rater who blindly evaluated the artifact-confounded data. To test the reliability of the procedure, another rater checked five randomly selected subjects using the same criteria. Inter-rater reliabilities were sufficiently high (Cronbach's $\alpha = 0.94$), and two subjects were eliminated due to the large quantity of artifacts produced (modified Thompson tau technique, $p < .01$; Wheeler & Ganji, 1996). For the data analyzed, there were few artifacts ($M = 3.6$, $SD = 3.1$), and no significant systematic differences between the four experimental conditions; Friedman's one-way analysis of variance, $\chi^2(3) = 1.05$, $p > .1$. The artifact-confounded data were rejected from the following analyses.

2.1.6.2. FACS coding for percent occurrence analysis. The subject's facial reaction in response to each stimulus was coded from the digitized videotape using Ekman and Friesen's (1978) FACS system. FACS is a comprehensive, anatomically based coding system that describes visible facial muscular movements in terms of AUs; coders do not make interpretations. Based on our research interests, only AU 4 (brow lowering, prototypical facial action in angry expressions) and AU 12 (lip corner pulling, prototypical facial action in happy expressions) were evaluated.

Two scorers trained in FACS coding blindly scored the subject's facial reaction within 2500 ms after stimulus onset. Inter-scorer reliabilities were sufficiently high (Cronbach's alpha = 0.90 and 0.96 for AU 4 and AU 12, respectively). The responses evaluated consistently by both scorers were subjected to the following analyses.

2.1.6.3. Statistical analysis of percent occurrence. Percent occurrence was calculated for each subject, for each AU. To satisfy normality assumptions for the subsequent analyses, these data were subjected to arcsine transformation (Cohen, Cohen, West, & Aiken, 2003). The figures, however, display the untransformed percentages.

The arcsined percent occurrences were analyzed with a 2 (presentation condition: dynamic/static) × 2 (expression: angry/happy) repeated measures design. Based on our specific prediction that facial expressions of emotion would elicit facial mimicry in the case of dynamic presentations but not in the case of static presentations, the interaction between the presentation condition and expression was analyzed. For the significant interactions, follow-up simple effect analyses were also conducted. These analyses were conducted using one-tailed *t*-statistics. The main effects of presentation condition and expression were also examined using two-tailed *t*-statistics, although we did not have specific hypotheses for these effects. Results of all tests were considered statistically significant at $p < .05$.

2.1.6.4. FACS coding for latency analysis. Two scorers trained in FACS coding evaluated the subjects' facial reactions frame by frame to describe the latency of the facial reactions. Based on the results of percent occurrence analyses, the AU 4 response to dynamic angry expressions and the AU 12 response to dynamic happy expressions, as identified through the above FACS coding, were analyzed. The latencies of these AUs were reliably identified by the scorers (Cronbach's $\alpha = 0.99$ and 0.94 for AU 4 and 12, respectively), and the mean values across the scorers were calculated. Because the dynamic facial expressions included the first neutral expressions for 240 ms after stimulus onset, the latency after the onset of dynamic changes was reported instead of the latency after stimulus onset.

2.2. Results

Representative examples of the subjects' facial reactions are presented as the static pictures in Fig. 2 and as the dynamic pictures on the following website: http://www.educ.kyoto-u.ac.jp/cogpsy/personal/yoshikawa/mimicry.html.

2.2.1. FACS data of percent occurrence

For AU 4 (Fig. 3, left), the interaction between presentation condition and expression was significant, $t(17) = 1.79$, $p < .05$. Simple effect analyses revealed that angry expressions elicited the AU 4 response in subjects more frequently than did happy expressions in the dynamic presentation condition, t $(17) = 1.99$, $p < .05$. The simple main effect of expression for static presentation was not significant, $p > .1$. The main effects of presentation condition and expression were also not significant, ps $> .1$.

For AU 12 (Fig. 3, right), the interaction between presentation condition and expression was significant, $t(17) = 2.14$, $p < .01$. Simple effect analyses indicated that happy expressions elicited AU 12 responses more frequently than did angry expressions in the dynamic presentation condition, $t(17) = 2.35$, $p < .01$. The simple main

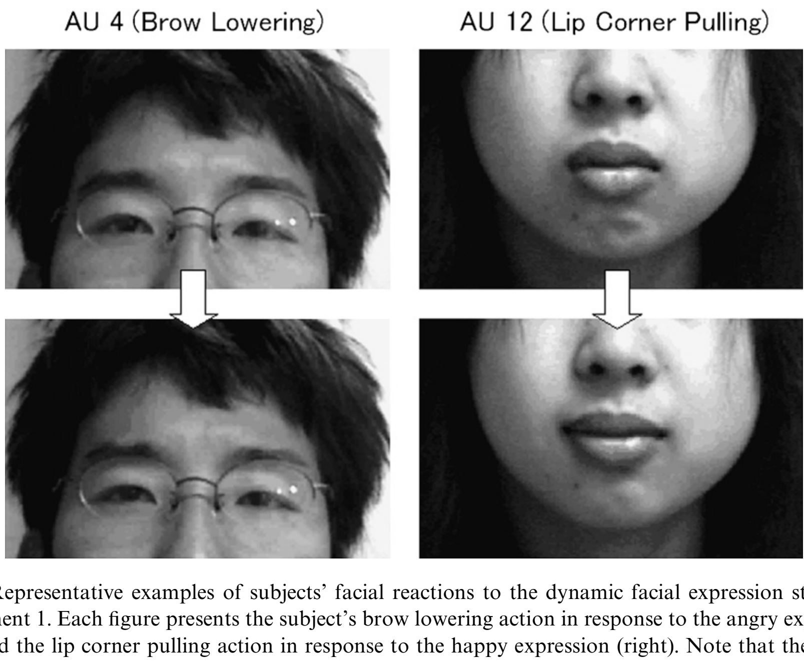

Fig. 2. Representative examples of subjects' facial reactions to the dynamic facial expression stimuli in Experiment 1. Each figure presents the subject's brow lowering action in response to the angry expression (left) and the lip corner pulling action in response to the happy expression (right). Note that the central portions of the brows were lowered and the brows were pulled together in the left figure, and the lip corners were elongated and angled up and cheeks were raised in the right figure.

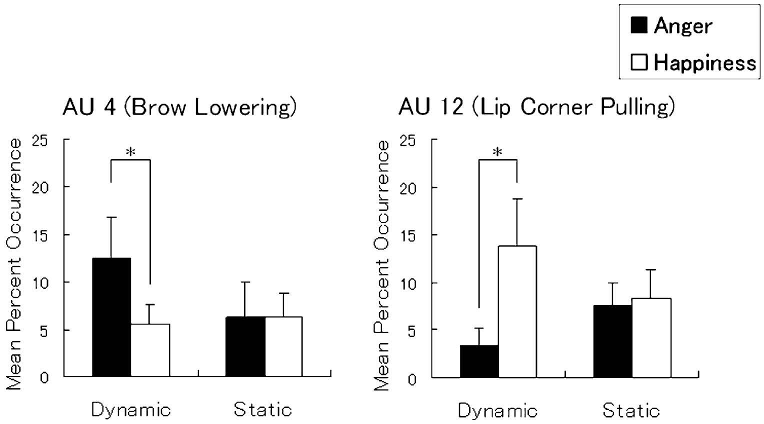

Fig. 3. Mean (with SE) percent occurrence of Action Unit (AU) 4 (left) and AU 12 (right) in Experiment 1.

effect of expression for static presentation was not significant, $p > .1$. The main effects of presentation condition and expression were not significant, $ps > .1$.

2.2.2. FACS latency data

For the AU4 response to dynamic angry expressions, the mean ($\pm SD$) latency from the onset of dynamic changes of facial expression was 874 (±262) ms.

For the AU12 response to dynamic happy expressions, the mean ($\pm SD$) latency from the onset of dynamic changes of facial expressions was 817 (±200) ms.

2.3. Discussion

The results of Experiment 1 revealed that specific facial action patterns were elicited in response to dynamic facial expressions of anger and happiness. Brow lowering actions, which are prototypical facial movements in angry expressions, occurred more frequently in response to angry expressions than to happy expressions, and lip corner pulling actions, which are prototypical facial actions in happy expressions, occurred more frequently in response to happy expressions than to angry expressions. Because the subjects passively viewed the stimuli, these facial actions reflected spontaneous processes. Latency analyses indicated that the facial responses occurred rapidly, specifically 800–900 ms after the onset of dynamic changes in the facial expressions viewed. These results support our hypothesis indicating that facial mimicry can be elicited spontaneously and rapidly in response to dynamic facial expressions of emotion.

The latency analysis showed that the mean latency of these actions was less than 900 ms after the onset of dynamic changes in facial expression. This result indicates that the facial reactions could have been elicited before the observation of final images under dynamic conditions that is, at this time the stimulus person and the subject showed simultaneous facial movement. This result is in line with previous studies investigating real life face-to-face communication (Condon & Ogston, 1967; Kendon, 1970). For example, Kendon (1970) analyzed conversational communication between individuals, and reported that listeners showed concurrent movement mirroring of speakers. Although people can use various types of context information in real communication (e.g., the stream of talk, Goodwin, 2002), our results indicate the synchronous mirroring facial movement between two individuals could occur by seeing the facial information of each other.

The rapid occurrence of facial responses is in line with previous EMG data (Dimberg & Thunberg, 1998), however, the specific latency was incongruent. Dimberg and Thunberg reported that the specific facial EMG difference occurred after 500 ms of exposure to the facial pictures. Because the elasticity of the facial sheath, facial skin, and adipose tissue attenuates the visible effects of rapid muscle contractions (Fridlund & Cacioppo, 1986), there may be a time lag between underlying muscle activity and the visible facial behaviors that require sufficiently intense muscle activity.

3. Experiment 2

Experiment 2 was conducted to investigate some issues that remained unresolved by Experiment 1. First, in Experiment 1 we used computer-morphing to present dynamic facial expressions. This method had the disadvantage of artificiality, and subjects' facial reactions may possibly have related to some confounding factors associated with this artificiality. For example, subjects may have smiled during dynamic happy expressions because the stimuli were unnatural. To examine this issue, we presented videos of real facial expressions in Experiment 2. This method has the advantage of natural and ecologically valid facial expressions, relative to computer-morphing; however, disadvantages of this method are that the facial expressions have idiosyncratic spatial and temporal parameters and often contain various artifacts irrelevant to emotional expressions, such as blinking and eye movements. Because it was hard to strictly control the temporal parameters of these stimuli, we did not analyze the latencies of the subjects' facial reactions. Second, because trained FACS coders evaluated the subjects' facial reactions in Experiment 1, it was still unclear whether normal viewers could discriminate the facial reactions. This point is crucial because if normal viewers did not recognize the facial reactions, the idea that the facial reactions contribute to intra-individual communication would not be warranted. To examine this issue, we asked naive raters to rate the subjects' facial reactions. We predicted that the specific facial reactions interpretable as facial mimicry could occur in dynamic presentations that even naive observers could recognize.

3.1. Methods

3.1.1. Subjects

Eighteen volunteers (14 females and 4 males; mean age, 19.2 years) participated in this experiment. All the subjects were right-handed, and had normal or corrected-to-normal visual acuity. Although three additional subjects actually participated, their data were not analyzed. One of them reported in debriefing that she had noticed being videotaped. The remaining two had hairstyles that hid their brows.

3.1.2. Experimental design

The design was identical to that used in Experiment 1.

3.1.3. Stimuli

Materials were video films of angry and happy facial expressions of eight females and eight males. These stimuli were selected from a video database of facial expressions of emotion posed by more than 50 Japanese models. None of the faces were familiar to any of the subjects. Preliminary ratings from subjects who did not take part in this experiment confirmed that the stimuli clearly displayed the target emotions relative to other basic emotions. We selected stimuli using the following criteria: the facial expressions contained valid spatial parameters, i.e., the stimuli showed AU 4 and 12 for angry and happy expressions, respectively; the expressions contained similar temporal parameters as the dynamic stimuli in Experiment 1; and the expressions contained few artifacts irrelevant to emotional expressions.

For the dynamic expression stimuli, a total of 38 frames from neutral to emotional expressions were presented. Each frame was presented for 40 ms, and each clip was presented for 1520 ms.

The frames of the apex emotional expressions in the dynamic condition were prepared for the static expression stimuli and presented for 1520 ms.

3.1.4. Apparatus

The apparatus was the same as in Experiment 1.

3.1.5. Procedure

The procedure was identical to that in Experiment 1 with three minor modifications. First, because some subjects in Experiment 1 reported that the stimulus presentations after the 520 ms fixation presentation period felt somewhat sudden, the fixation point was extended to 1520 ms.

Second, for the rating after viewing all the images, a dummy questionnaire was prepared. The subjects were asked to rate the stimuli using the questionnaire. These data are not reported here because the performance of this task was irrelevant to the purposes of the study.

Third, debriefing was conducted immediately after the experiment, and subjects' permission to use the videotapes for rating by naive raters was additionally requested. Permission was granted in all cases.

3.1.6. Data analysis

3.1.6.1. Preprocessing. The method of artifact rejection was identical to that used in Experiment 1. Inter-rater reliabilities for five randomly selected data were sufficiently high (Cronbach's $\alpha = 0.87$). None of the subjects was eliminated owing to the artifacts (modified Thompson tau technique, $p > .1$; Wheeler & Ganji, 1996). The analyzed data had few artifacts ($M = 3.4$, $SD = 2.8$). There were also no significant systematic differences among the four experimental conditions (Friedman's one-way analysis of variance, $\chi^2(3) = 4.29$, $p > .1$). Artifact-confounded data were rejected from the following analyses.

3.1.6.2. FACS coding for percent occurrence analysis. The FACS coding method was identical to that used in Experiment 1. Inter-scorer reliabilities in this experiment were again sufficiently high (Cronbach's $\alpha = 0.91$ and 0.94 for AU 4 and AU12, respectively).

3.1.6.3. Recognition by naive observers. Thirteen volunteers (3 females and 10 males; mean age, 25.2 years) participated in the recognition of the subjects' video data. The raters were different from those who took part in the video recording and FACS coding. They were blind to the conditions and hypotheses of the experiment. Each participated in this experiment in an individual chamber room.

We randomly selected eight video data from each of following four conditions: data for dynamic anger with and without AU 4, and data for dynamic happiness with and without AU 12. These stimuli were individually presented on a 19-in. liquid crystal monitor (L767, FlexScan) using a Windows-compatible computer (PCV RZ752, Sony). The order of the presentations was randomized.

The raters were first exposed to all the data. Then, they were asked to rate whether facial expressions of negative or positive emotion were displayed; raters had three choices: negative, positive, or neutral expressions. It was emphasized that the face in the first frame of each clip was the reference facial expression for the stimulus person and that change from that face should be evaluated. Raters were allowed to repeatedly watch the video data.

In the data analysis, percentages of the negative, positive, and neutral expression recognition were calculated for each rater. To satisfy normality assumptions for the subsequent analyses, these data were subjected to arcsine transformation (Cohen et al., 2003). Then, the following two analyses were conducted.

First, the percentage of negative expression recognition in response to the data for dynamic anger with AU 4, and the percentage of positive expression recognition in response to the data for dynamic happiness with AU 12 were tested for difference from chance using one-sample t tests (one-tailed, $p < .05$).

Second, recognition percentages of negative and positive expressions across the four conditions were analyzed using multiple comparisons by Dunnet methods (one-tailed, $p < .05$).

3.2. Results

Representative examples of the subjects' facial reactions are presented as the static pictures in Fig. 4 and as the dynamic pictures on the following website: http://www.educ.kyoto-u.ac.jp/cogpsy/personal/yoshikawa/mimicry.html.

3.2.1. FACS data of percent occurrence

For AU 4 (Fig. 5, left), the interaction between presentation condition and expression was significant, $t(17) = 1.79$, $p < .05$. Simple effect analyses revealed that angry expressions elicited subjects' AU 4 response more frequently than did happy expressions in the dynamic presentation condition, $t(17) = 2.37$, $p < .05$. The simple main effect of expression for static presentation was not significant, $p > .1$. The main effect

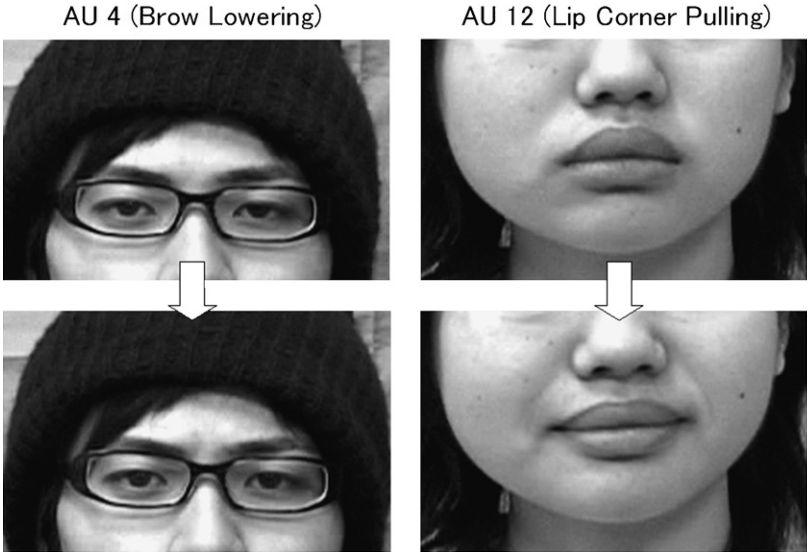

Fig. 4. Representative examples of subjects' facial reactions to the dynamic facial expression stimuli in Experiment 2. Each figure presents the subject's brow lowering action in response to the angry expression (left) and the lip corner pulling action in response to the happy expression (right). Note that the inner portions of the brows were lowered and the brows were pulled together in the left figure, and the lip corners were elongated and angled up and cheeks were raised in the right figure.

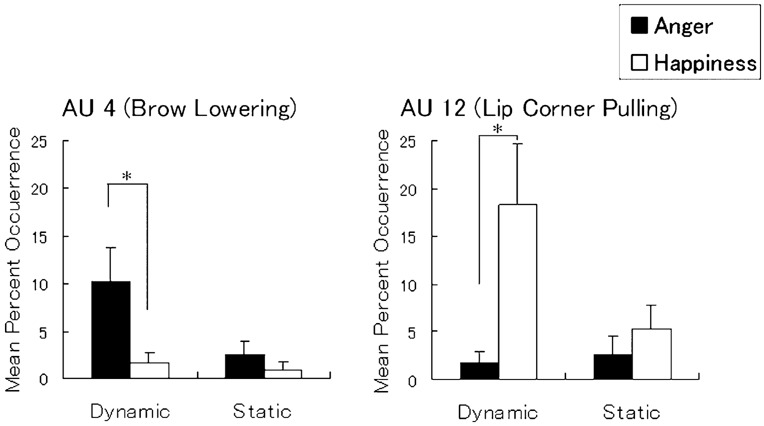

Fig. 5. Mean percent occurrence of Action Unit (AU) 4 (left) and AU 12 (right) in Experiment 2.

of expression for the same pattern was also significant, $t(17) = 2.49$, $p < .05$, indicating more frequent AU 4 for angry than for happy expressions. The main effects of presentation condition reached marginal significance, $t(17) = 2.02$, $p < .1$.

For AU 12 (Fig. 5, right), the interaction between presentation condition and expression was significant, $t(17) = 2.87$, $p < .01$. Simple effect analyses indicated that happy expressions elicited the AU 12 response more frequently than angry expressions under the dynamic presentation condition, $t(17) = 3.14$, $p < .005$. The simple main effect of expression for static presentation was not significant, $p > .1$. The main effect of expression was significant, $t(17) = 3.14$, $p < .01$, indicating more frequent AU 12 for happy than angry expressions. The main effect of presentation condition was also significant, $t(17) = 3.29$, $p < .005$, indicating more frequent AU 12 for dynamic than for static presentations.

3.2.2. Recognition by naive raters

For the percent recognition of negative expression in response to the data for dynamic anger with AU 4 (Fig. 6, left), the one-sample *t* test showed that the percent significantly differed from chance, $t(12) = 16.77$, $p < 001$. Multiple comparisons using Dunnet methods revealed that the data for dynamic anger with AU 4 were recognized as negative more frequently than for all other conditions, *p*s < .05.

For the percent recognition of positive expression in response to the data for dynamic happiness with AU 12 (Fig. 6, right), the one-sample *t* test showed that the percentage significantly differed from chance, $t(12) = 12.05$, $p < 001$. Multiple comparisons using Dunnet methods revealed that the data for dynamic happiness with AU 12 were recognized as positive more frequently than all of the other conditions, *p*s < .05.

3.3. Discussion

The results of FACS coding revealed that specific facial action patterns interpretable as facial mimicry were elicited in response to dynamic facial expressions of anger and happiness, as in Experiment 1. Because we used the videos of natural changes in facial expressions under dynamic conditions, these results could not be attributable to artifacts related to artificial dynamic stimuli. The results also revealed that the naive raters recognized the subjects' facial reactions as emotional expressions in which valence corresponded to the dynamic facial expressions the subjects were viewing. These results support our hypothesis that facial mimicry can be elicited

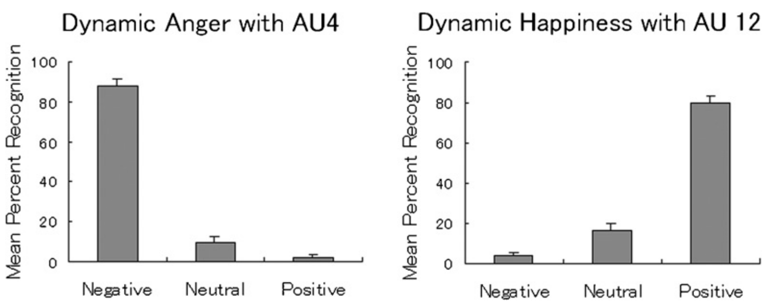

Fig. 6. Mean (with SE) percent recognition of each emotional expression (negative, neutral, or positive) by naive raters in Experiment 2. Each figure depicts the results in response to the video data for dynamic anger with Action Unit (AU) 4 (left) and dynamic happiness with AU 12 (right).

spontaneously in an externally visible form in response to the dynamic facial expressions of emotion.

In contrast to the results of Experiment 1, the main effect of expression reached significance both in AU 4 and 12. Although the interpretation of main effects could be problematic when interactions were significant (Tabachnick & Fidell, 2001), visual inspection of the figures suggests that the patterns of the effect of expression were similar across dynamic and static presentation conditions. The use of eight people for the stimuli in Experiment 2, instead of four people as in Experiment 1, may have reduced the saturation to static stimuli. Alternatively, because we used video for the dynamic conditions in this experiment, it is possible that this context familiar to daily life enhanced the facial reactions in response to static facial expressions. Previous facial EMG studies (e.g., Dimberg, 1982) have consistently indicated facial muscle activity in response to static facial expressions; therefore, we do not believe that the static facial expressions had no capacity to induce spontaneous facial mimicry. Our results only indicate the relative advantage of dynamic presentations compared to static presentations in the elicitation of facial mimicry.

4. General discussion

The results of the two experiments consistently indicated that specific facial action patterns, interpretable as facial mimicry, were spontaneously elicited in response to dynamic facial expressions of anger and happiness. By using two different methods to present dynamic facial expressions, (i.e., the computer-morphing techniques in Experiment 1 and videos of natural facial expressions in Experiment 2), the equivalent occurrence of facial reactions was confirmed. Additionally, latency analyses in Experiment 1 indicated that the facial responses occurred rapidly, specifically 800–900ms after the onset of the dynamic changes in the facial expressions viewed. Analysis of recognition by naive raters in Experiment 2 indicated that ordinary people recognized the subjects' facial reactions as emotional expressions in which the valence corresponded to the dynamic facial expressions the subjects were viewing. These results are in line with previous studies, which recorded facial EMG and found that the presentation of pictures of facial expressions elicited spontaneous and rapid facial muscular activity, interpretable as facial mimicry (e.g., Dimberg & Thunberg, 1998). However, muscle activity measured in those studies was too subtle, and it was unclear as to whether this activity could be visibly perceived as facial mimicry. The present results are also consistent with previous developmental studies showing that newborn infants imitate the facial gestures of adults (e.g., Field et al., 1982); our results provide similar findings for adult subjects. Taken together, the results of our study indicate that facial mimicry can be elicited spontaneously and rapidly in an externally visible form in response to dynamic facial expressions of emotion.

Facial mimicry was more evident for dynamic presentations relative to the static presentations. This result is consistent with those of a neuroimaging study (Sato et al., 2004) indicating that the inferior frontal gyrus (BA 44) was more evidently active in response to dynamic facial expressions than to static facial expressions. The homologue ventral premotor cortex of monkeys has been shown to contain mirror neurons that discharge while both performing specific actions and observing another individual performing similar actions (Rizzolatti & Arbib, 1998). We speculate that the mirror neuron system may play an important role in facial mimicry, eliciting a subject's facial reactions to another individual's facial movements. Visible facial behaviors that follow those of other individuals probably facilitate inter-individual communication and social coordination; this neuro-cognitive mechanism may be of critically adaptive significance for social primates, including humans.

Although distinct facial patterns were elicited while viewing dynamic facial expressions, their frequency was low. Both brow lowering and lip corner pulling occurred less than 20% of the time. Comparable data were reported in a previous study (Field et al., 1982), which found that infants showed lip widening in response to an adult model's expressions of happiness at a frequency of approximately 12%. It seems that facial mimicry in response to the facial expressions of other individuals occurs at a relatively low frequency and it does not occur regularly. Based on a series of infant studies, Meltzoff and colleagues (e.g., Meltzoff, 2005) proposed that infant facial mimicry consists not only of reflexive components but also active execution/inhibition processes. It is plausible that adults also have the mental mechanism to inhibit spontaneous facial mimicry. Such a mechanism would help control face-to-face communication from being too emotionally resonant. It may be possible that facial mimicry only occasionally overcomes inhibition.

Although we specifically analyzed AU 4 and 12, according to our hypothesis of facial mimicry, our preliminary analyses indicated that the subjects showed some other types of facial actions. For example, entire brow raising (AU 1 + 2), which is a prototypical facial action in a surprised facial expression, occurred for dynamic angry expressions. Although these facial reactions were relatively infrequent and differed idiosyncratically across subjects, further investigation of non-mimicking facial reactions is an important matter for future research.

Promising directions for further investigation include discovering in more detail the factors that influence the quality and the quantity of spontaneous facial mimicry. Although we used faces unfamiliar to our subjects, familiar faces, for example, may elicit stronger facial movements. Although a passive viewing situation was used in the present study, active processing, such as sympathetic concern, may also increase facial mimicry.

In summary, we found that the overt specific facial actions, interpretable as facial mimicry, were elicited spontaneously and rapidly when perceiving dynamic facial expressions of anger and happiness. This clearly suggests that facial mimicry is not only a form of intra-individual processing, but that it also functions as a form of inter-individual communication. The results also support the importance of using dynamic stimuli to reveal the psychological mechanisms of real-life facial expression processing.

Acknowledgements

The authors heartily thank Dr. Yumi Fujitsu for her technical support, and Professor Jacques Mehler and anonymous reviewers for their helpful advice. This study was supported by Special Coordination Funds for promoting Science and Technology from The Science and Technology Agency of the Japanese government, and by a Research Fellowship from the Japan Society for the Promotion of Science.

References

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: an fMRI study. *Journal of Cognitive Neuroscience, 16*, 114–126.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology, 50*, 260–268.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research, 5*, 221–235.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology, 19*, 643–647.

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology, 39*, 39–45.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*, 86–89.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologist.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17*, 1703–1714.

Field, T., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitalion of facial expressions by neonates. *Science, 218*, 179–181.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*, 567–589.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

Goodwin, C. (2002). Time in action. *Current Anthropology, 43*, 19–35.

Hess, U., Philippot, P., & Blairy, S. (1999). Mimicry: Facts and fiction. In P. Philippot, R. Feldman, & E. Coats (Eds.), *The social context of nonverbal behavior* (pp. 213–241). New York: Cambridge University Press.

Keltner, D., & Kring, A. (1998). Emotion, social function, and psychopathology. *Review of General Psychology, 2*, 320–342.

Kendon, A. (1970). Movement coordination in social interaction: some examples described. *Acta Psychologica, 32*, 101–125.

McHugo, G. J., Lanzetta, J. T., Sullivan, D. G., Masters, R. D., & Englis, B. G. (1985). Emotional reactions to a political leader's expressive displays. *Journal of Personality and Social Psychology, 49*, 1513–1529.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75–78.

Meltzoff, A. N. (2005). Imitation and other minds: the "like me" hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 55–77). Cambridge: MIT Press.

Mukaida, S., Kamachi, M., Kato, T., Oda, M., Yoshikawa, S., & Akamatsu, S. (2000). *Foolproof utilities for facial image manipulation*. Kyoto: ATR.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends Neurosciences, 21*, 188–194.
Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Research. Cognitive Brain Research, 3*, 131–141.
Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Brain Research. Cognitive Brain Research, 20*, 81–91.
Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition and Emotion, 18*, 701–710.
Tabachnick, B. G., & Fidell, L. S. (2001). *Computer-assisted research design and analysis*. Boston: Allyn and Bacon.
Wheeler, A. J., & Ganji, A. R. (1996). *Introduction to engineering experimentation*. Englewood Cliffs, NJ: Prentice Hall.