# Facial mimicry in the congenitally blind

Pablo Arias, Caren Bellmann, Jean-Julien Aucouturier

# Facial mimicry in the congenitally blind

Pablo Arias[1,2,*], Caren Bellmann[3] and Jean-Julien Aucouturier[1,4]

*[1]STMS Lab (IRCAM/CNRS/Sorbonne Université), 1 Place Igor Stravinsky, 75004 Paris, France [2]Lund University Cognitive Science, Lund University, Box 192, 221 00 Lund, Sweden. [3]Institut National des Jeunes Aveugles, 56 Bd des Invalides, 75007 Paris, France. [4]FEMTO-ST Institute (CNRS/Université Bourgogne Franche Comté), 15B Av. des Montboucons, 25000 Besançon, France *Lead contact.*

Imitation is one of the core building blocks of human social cognition, supporting capacities as diverse as empathy, social learning, and knowledge acquisition[1]. Newborns' ability to match others' motor acts, while quite limited initially, drastically improves during the first months of development[2]. Of notable importance to human sociality is our tendency to rapidly mimic facial expressions of emotion. Facial mimicry develops around six months of age[3], but because of its late emergence, the factors supporting its development are relatively unknown. One possibility is that the development of facial mimicry depends on seeing emotional imitative behavior in others[4]. Alternatively, the drive to imitate facial expressions of emotion may be independent of visual learning and be supported by modality-general processes. Here we report evidence for the latter, by showing that congenitally blind participants facially imitate smiles heard in speech, despite having never seen a facial expression.

To investigate whether facial mimicry develops independently from visual learning, we studied how blind participants respond to the acoustic cues generated by a smiling facial expression while speaking[5]. To control these cues in experimental stimuli, we used a digital audio processing algorithm that simulates how the contraction of zygomatics shifts spectral resonances — formants — in the voice[6] (Figure 1A), while leaving all other characteristics of emotional speech, such as content, or intonation, unchanged. Using this tool, we generated 120 spoken-sentence stimuli, by transforming 40 sentences in three matched conditions: neutral, smile (increased lip stretching) and unsmile (decreased lip stretching). In these stimuli, the transformation had the notable effect of selectively shifting the mean frequency of the first two

vocal formants either positively (smile effect) or negatively ('unsmile' effect; $p < 0.0001$, Figure 1B; Supplemental Information).

Using these stimuli, we conducted an electromyography (EMG) experiment to study facial mimicry in the blind. We asked N = 14 blind participants — five congenital, six early, three late; all purely ocular, non-cortical impairments — to judge the smiliness of the generated stimuli in two successive tasks: a rating task (continuous rating scale) and a detection task (go/no go). In both tasks, participants rated the smiled and unsmiled versions of all sentences, while we recorded their zygomatic major (used to smile) and corrugator supercili (used to frown) muscles with facial EMG (see Supplemental Information for detailed experimental procedures).

As a manipulation check, the acoustic manipulation significantly affected participants' impression of speaker's smiliness both in the rating ($\chi^2(11) = 16.46$, $p = 0.0003$) and in the detection task ($\chi^2(5) = 35.1$, $p = 2.38 \times 10e^{-8}$; Figure 1C; Supplemental Information). Individual statistics confirmed that blind participants significantly recognised the auditory signature of smiles in stimuli (congenital: 4/5, 80%; early: 6/6, 100%; late: 1/3, 33%; all: 11/14, 79%; Supplemental Information). Smile-detection accuracy was comparable with that of previously tested[6] sighted controls (Welsh's unequal variance t-test $t(13.0) = 1.95$, $p = 0.07$, Figure 1D, Supplemental Information).

We then analysed the difference between smile and unsmile EMG activity with Generalized Linear Mixed Models (GLMMs), combining data from both tasks, and found clear evidence of facial mimicry at the group level across all blind participants. For the zygomatic muscle, we found a main effect of the sound manipulation ($\chi^2(1) = 4.56$, $p = 0.03$). The smile manipulation significantly increased zygomatic activity by 1.14 (±0.5 SE, $p = 0.03$) when compared to the unsmile effect. Conversely, for the corrugator muscle, the smile manipulation decreased muscle activity, although the difference was not significant ($\chi^2(1) = 1.4$, $p=0.24$; see Supplemental Information for in-depth analysis of each task; Supplemental Data S1E).

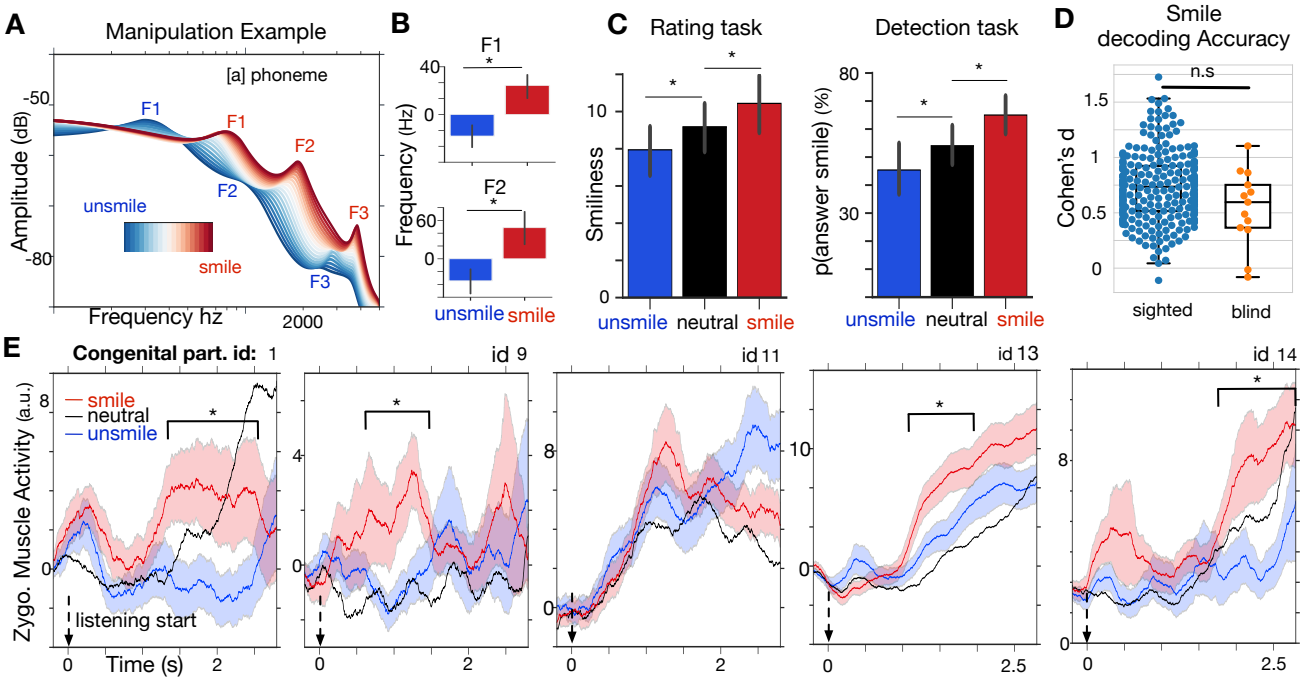



Figure 1. **Controlling and perceiving auditory smiles.** (A) Audio manipulation example of an [a] phoneme, where can be seen the formant movements from the unsmile (blue) to the smile transformation (red) (B) Formant analysis of the stimuli for both unsmile (blue) and (smile) manipulations; Formants were normalised by the non-manipulated (neutral) sound; asterisks indicate significant differences between the distributions; error bars are 95% confidence intervals. (C) Mean smiliness rating (left) and decoding accuracy (right) for unsmile, neutral and smile transformations. (D) Difference in smile detection accuracy between sighted and blind participants. Each point represents the Cohen's d for an individual participant, computed using smile and unsmile rating distributions (rating task). Sighted participant data were simulated using a previously collected dataset[6], see Supplemental Information. Welsh's unequal variance t-test (13.0) =1.95, p = 0.07; n.s., not statistically significant (see Supplemental Information). (E) Zygomatic activity for congenitally blind participants during the listening of the stimuli for smile (red), neutral (black) and unsmile (blue) conditions; Shaded areas represent SEM; Asterisks indicate significant differences between smile and unsmile time series ($p < 0.05$)

To investigate the case of congenital participants specifically, we then analysed the EMG time series with individual statistics. We used the 240 time-series for each participant and each muscle and cluster permutation tests[7]. We found 10 clusters differentiating smile and unsmile EMG time series, all of which were congruent with the acoustic manipulation (four for the zygomatic muscle; six for the corrugator muscle, $p < 0.05$; Figure S1 and Figure S2 in the Supplemental Information; Supplemental Data S1F). Across the blind group, both the number of significant clusters, and their effect sizes, did not differ from sighted controls (Supplemental

Information; Supplemental Data S1G). Crucially, significant clusters of congruent muscle activity were present in ⅘ (80%) of our congenitally blind participants (Figure 1E).

In sum, we present here robust, replicated evidence that congenitally blind individuals are not only able to recognise smiling speakers from the sound of their voice, but also to implicitly mirror these smiles in their own facial expression in a similar manner to sighted individuals. The fact that our participants recognized auditory smiles is in contrast with the fact that blind individuals generally have difficulty recognizing emotions from vocal tones[8]. This suggests that, contrary to prosody, learning 'how smiles sound' does not heavily rely on the availability of contextual information about the faces of one's conversation partners, perhaps because their acoustic signature affords more direct inferences about a speaker's oro-facial configuration than does a given contour of pitch or loudness[6].

More importantly, while it is known that congenitally blind individuals have preserved abilities to produce smiles and other facial expressions of emotions[8], the fact that they do so spontaneously in response to auditory smiles constitutes striking evidence of facial mimicry in participants who, yet, have never seen a facial expression. While there is debate on whether facial imitative behavior develops on the basis of learned or innate associations[2], most theories of imitation place visual observation as a core building block of imitative mechanisms[4,9]. Here, the fact that congenitally blind participants imitate smiles heard in speech conclusively demonstrates that the mechanisms of facial mimicry in fact do not require visual learning to develop.

How, then, did this capacity emerge? Consistent with the associative learning view[4], it is possible that, for blind individuals, auditory-motor associations heard in vocalizations and experienced in one's own proprioception provide a non-visual route for learning to perceive and produce facial expressions of emotion. In the alternative innate view, it is also possible that these associations do not require learning and are built in the system, either in the form of cortical mirror mechanisms[9] or of prewired emotional responses taking input from phylogenetically-ancient, multimodal (visual-auditor-motor) subcortical structures[10].

In either case, the present results demonstrate that imitation is not a mere visuo-motor process, but rather a flexible mechanism deployed across sensory inputs, able to map cross-modal exteroceptive signals to their corresponding motor representations and socially-appropriate responses.

## Acknowledgments

The authors thank Lou Seropian, Pilar Rios, Isabelle Perrault and Josseline Kaplan for help with participant recruitment and data collection. Experimental data collected at INJA (National Institute for Blind Youth) with support from the INSEAD/Sorbonne University Centre for Behavioural Science. Work funded by ERC StG CREAM 335536, ANR REFLETS and SEPIA, as well as support from Fondation pour l'Audition FPA RD-2018-2 (to JJ.A.), and the Swedish Research Council (2014-1371) to P.A.

## References

1. Frith, C.D., and Frith, U. (2012). Mechanisms of social cognition. Annu. Rev. Psychol. *63*, 287-313.

2. Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., Clark, S., and Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. Curr. Biol. *26*, 1334-1338.

3. Kaiser, J., Crespo-Llado, M.M., Turati, C., and Geangu, E. (2017). The development of spontaneous facial responses to others' emotions in infancy: An EMG study. Sci. Rep. *7*, 17500.

4. Heyes, C. (2016). Homo imitans? Seven reasons why imitation couldn't possibly be associative. Phil. Trans. R. Soc. Lond. B *371*, 20150069.

5. Tartter, V.C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. Attent. Percept. Psychophys. *27*, 24-27.

6. Arias, P., Belin, P., and Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. Curr. Biol. *28*, R782--R783.

7. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Meth. *164*, 177-190.

8.Valente, D., Theurel, A., and Gentaz, E. (2018). The role of visual experience in the production of emotional facial expressions by blind people: a review. Psychon. Bull. Rev. *25*, 766–774.

9. Meltzoff, A.N., and Marshall, P.J. (2018). Human infant imitation as a social survival circuit. Curr. Opin. Behav. Sci. *24*, 130–136.

10. Johnson, M.H. (2005). Subcortical face processing. Nat. Rev. Neurosci. *6*, 766–774.

# Supplemental Information: Facial Mimicry in the Congenitally Blind

Pablo Arias, Caren Bellmann, Jean-Julien Aucouturier

## Supplemental Experimental Procedures

### Methods

**Participants:** N=14 right-handed, French-speaking, blind participants (female:5, male:9, Mean age=33.5, min=21, max=58) took part in this experiment. We divided participants in three groups[S1]: 5 congenital (never had sight), 6 early (lost their sight before being 13 years of age) and 3 late participants (lost their sight after being 13 years of age).

Participants were included on the basis of prior medical screening by the second author, consultant ophthalmologist at the National Institute for Blind Youth (Institut National des Jeunes Aveugles, INJA) in Paris, who confirmed that all participants' blindness was due to ocular rather than neurological factors (i.e. no participant had blind sight) and that no participant had psychiatric or neurological conditions that could interact with the task (such as autism spectrum disorder, a frequent comorbidity with visual impairment[S2]). In addition, participants reported having no hearing impairments. Participant 6 was excluded from all EMG analysis because of technical problems during the experiment.

Across the group, visual acuity was limited to light perception; one participant had vague movement perception. 6 participants had complete blindness, 7 subjects had light perception (WHO category 4 and 5) and one subject (id: 5, category late) could vaguely perceive hand movements. None of the participants had vision sufficient to identify facial expressions. Data S1A presents a complete etiology of participants' blindness.

**Stimuli:** The smile gesture is thought to alter formant frequencies in speech[S3]. We developed a digital audio algorithm capable of recreating these acoustic changes in running speech (as if the speech was produced with/without smiles) but leaving all other aspects of emotional speech unchanged (i.e. semantic content, pitch contour, speech rate, length, temporal dynamics, and speaker gender are kept constant across conditions). In short, the smile transformation increases the first two formant frequencies in speech, whereas the unsmile transformation implements the opposite acoustic transformation (decreasing formant frequencies). Previous work describes the technical details of the digital audio algorithm[S4] and its experimental validations[S5]. Sound examples can be found in previous publications[S4,S5].

We used this digital audio algorithm to create the stimuli for the present study. 40 sentences were recorded by male and female native French speakers, and transformed using the smile and unsmile transformations, resulting in 40 neutral, 40 smile- and 40 unsmile-transformed sounds, for a total

of 120 stimuli. Mean stimulus duration was 1.9s seconds (SD=1.4s). All stimuli were normalised in loudness at 70 dbA using a Matlab toolbox[S6].

In previous work we have shown that the smile audio transformation influences emotion perception[S5], but it doesn't simply map to happy/sad expressions. Indeed, the verbal content and prosody of the original sentence are kept unchanged by the transformation, and these are important factors in shaping emotional judgements. For instance, sentences transformed with the smile effect can be perceived as more joyful, but also more ironic[S5]. Note that the current study contrasts ratings and EMG reactions to pairs of matched stimuli, composed of the same sentence modified with both the smile and unsmile transformation. This procedure therefore cancels out the effect of sentence prosody, verbal content, speaker identity or sex.

**Acoustic Analysis:** To ensure that the algorithmic manipulation shifts formant frequencies in running speech, we computed the stimuli's first and second formant frequencies (F1 and F2) using the Praat software[S7]. We normalised formant frequencies by their corresponding non-manipulated values and then compared formant distributions by means of paired t-tests. We found that the smile effect significantly increased formant frequencies for both F1 ($t(39)=5.63$, $p=1.6e-6$) and F2 ($t(39)=4.24$, $p=0.0001$) as compared to unsmile transformations. Data is presented in main Figure 1-B.

**Procedure:** Participants began by sitting in a chair in the experimental room. The experimenter explained that the aim of the study was to investigate how blind individuals perceive emotional speech. After cleaning participants' face with an antiseptic solution, the EMG sensors were placed in the zygomatic major and the corrugator supercili muscles[S8]. In order to divert participants from the true purpose of the experiment, the experimenter told a cover story stating that the EMG sensors were sweat sensors. The true aim of the physiological recordings was explained immediately after the experiment, during the debriefing session.

The experiment consisted of two separate blocks. In the first block, participants were presented with the 120 audio stimuli using a Beyerdynamic DT-770 headphones and an audio interface (RME Fireface UCX). Stimuli were pseudo-randomised by maximizing the distance of presentation of sentences from the same sound token. During the first task, participants were asked to answer for each stimulus "to what extent [was] this sentence pronounced with a smile" using a unipolar continuous scale ranging from 0 ("not smiling") to 20 ("a lot of smile") (subsequently called 'rating-scale block').

In block two, participants were presented with the same 120 stimuli as in block one, in a new pseudo-random order, but were asked this time to choose, for each sentence, whether the sentence was pronounced with or without a smile in a "go/no go" task, which was followed by a confidence judgement rating, ranging from 1 to 4 (1 : "I am not sure I gave the correct answer"; 4 : "I am sure I gave the correct answer"). This block is subsequently called the 'detection block'.

The experiment was coded in python using the open source software psychopy[S9]. All ratings were performed with the computer keyboard. Instructions were given at first by the experimenter and then by a vocal synthesizer based on apple's built-in speech synthesis engine controlled in real time by a python wrapper. At all times (except during the listening of the stimuli) participants

could interact with the vocal synthesizer by pressing keyboard commands to hear the instructions, the questions or the labels of the scales.

**EMG apparatus and pre-processing:** Electromyography (EMG) activity from corrugator supercili and zygomaticus major muscles was recorded during the listening of the stimuli on the left side of the face at Fs = 1000 Hz. Three online filters were used during the EMG recording: a high-pass filter at 10 Hz, a notch filter at 50 Hz and a low-pass filter at 499 Hz. EMG activity was recorded using two bipolar montages (BIP2AUX adapter), an ActiChamp amplifier, and Brainvision recorder software. Synchronization between the stimuli and the recording computer was done via the Cedrus StimTracker serial port. Offline, data was filtered with a 50Hz high-pass IIR filter and a 250Hz low-pass IIR filter, then segmented into 3.8s epochs (which include 800ms pre-stimulus baseline). Epochs were rectified and smoothed using a moving average function with a window of 300 ms, and finally z-score normalised with respect to each trial's baseline.

Artifact rejection was performed by computing a rejection threshold[S10]. We first excluded all trials where the absolute mean EMG activity was above 100 times the baseline's activity, and then fixed the artifact rejection threshold as three times the STD of the resulting distribution (in our data, 48 times the baseline activity). In total, there were 6240 EMG recordings (13 participants (14-1) x 2 blocks x 120 sounds x 2 muscles), from which we discarded 117 trials, which represent 1.8% of the total number of trials of the dataset.

**Ethics:** All experiments were approved by the Institut Européen d'Administration des Affaires (INSEAD) IRB. In accordance with the American Psychological Association Ethical Guidelines, all participants gave their informed consent and were debriefed and informed about the purpose of the research after the experiment.

## Ratings Data Analysis

### Group analyses

Participant ratings in both tasks were analysed using Generalized Linear Mixed Models (GLMMs). In all the following analyses (except cluster permutation tests), we report p-values, estimated from hierarchical model comparisons using likelihood ratio tests[S11], and only present models that satisfy (1) the assumption of normality (validated by visually inspecting the plots of residuals against fitted values), and (2) statistical validation (significant difference with the nested null model). To test for main effects, we compared models with and without the fixed effect of interest. To test for interactions, we compared models including fixed effects versus models including fixed effects and their interaction.

In the rating-scale task, we found a significant main effect of the sound transformation (3 levels: unsmile, neutral, smile; $\chi_2(11) = 16.46$, p=0.0003; main Figure 1-C). For the model including sound transformation as a predictor, the unsmile effect significantly lowered the smile ratings from the non-modified (neutral) sound by about -1.24 ± 0.30 (standard errors; p=7.85e-05; d=-0.50). Conversely, the smile effect significantly increased the smile ratings, by about 1.26 ± 0.41 (standard errors; p=0.0089; d=0.47) when compared to the neutral sound. File token and participant number were used as random factors in the GLMM.

In the detection task, we computed the probability of answering "smile" for each transformation category (3 levels: unsmile, neutral, smile) for each participant (main Figure 1-C). We performed a similar GLMM analysis as for the rating task but using only "participant number" as random factor. Indeed, as we used the overall discrete ratings (either "unsmile" or "smile") to compute participants' overall detection rate, "sound token" could not be used as a random factor. As for the rating task, we found a significant main effect of the sound transformation in the detection task ($\chi_2(5)$ =35.1, p=2.38e-08; main Figure 1-C). For the model including sound transformation as a predictor, the unsmile effect significantly lowered the smile ratings from the non-modified (neutral) sound by about -0.087 ± 0.02 (standard errors; p=0.0008; d=-0.55). Conversely, the smile effect significantly increased the smile ratings, by about 0.11 ± 0.02 (standard errors; p=5.61e-05; d=0.82) when compared to the neutral sound.

**Relation between blindness onset and decoding accuracy**

In order to examine individual differences of ratings within the group, we computed the difference between each participant's ratings of the smile and unsmile effect in the rating-scale task, and correlated it with participant's onset of blindness (Supplemental Data S1B). There was an apparent negative relation between the rating sensitivity to the effect and the onset of blindness, although the correlation was not statistically significant (p=0.09; r=-0.5).

**Statistics at the individual level**

To examine whether individual participants significantly recognised the acoustic signature of smiling, we performed individual statistics for both the rating and the detection tasks using GLMMs.

For the rating task, we fitted participants' continuous ratings with a model containing sound transformation (3 levels: unsmile, neutral, smile) as a predictor and sound token as a random factor. For each participant, we compared this model to the nested null model with likelihood ratio tests. Results are presented in Supplemental Data S1C.

We performed a similar analysis with the data from the detection block. We fitted participant's discrete ratings (2 levels: smile vs unsmile) with a model containing sound transformation as a predictor (3 levels: unsmile, neutral, smile) and sound token as a random factor. Results are presented in Supplemental Data S1C.

Auditory smiles' recognition rate for the different blind categories were as follows: 4/5 (80%) congenital participants, 6/6 (100%) early participants and 1/3 (33%) late participants showed a significant recognition of the signature of smiles in at least one of the two tasks.

## EMG data – Group analyses

**Group analyses – Rating Task**

To compare smile and unsmile EMG time series at the group level we performed cluster permutation analysis[S12]. For each muscle and for each participant, we computed the mean EMG time series for both smile and unsmile conditions. Cluster permutation tests revealed that the smile effect had a significant effect on EMG activity in the rating task, where zygomaticus major activity was congruent with the acoustic manipulation, that is, higher in the smile condition as compared to the unsmile condition (p-value: 0.04; time: 2.0-2.7; peak: 2.4).

To understand the link between explicit ratings and facial reactions, we collapsed EMG measures across participants, for each stimuli and each muscle, using the data from the rating task. To do this, we computed the mean across the time axis by taking EMG time series from 0 to 3 seconds (during the listening of the stimuli). This way, we computed 120 data points distributed along the 'rating' axis. We then correlated the mean rating of these 120 data points with their mean EMG activity, for each muscle. Note that each data point represents the mean muscle activity triggered by a specific stimulus across participants. Data is presented in Supplemental Data S1D.
For the zygomatic muscle we found a significant correlation between muscle activity and rating ($p=0.03$, $r=0.19$). We found no significant correlation for the corrugator muscle ($p=0.21$, $r=-0.11$). We then performed the same analysis but grouping the data between congenital, early and blind groups. We found a significant correlation for the congenital participants for the zygomatic muscle ($p=0.03$, $r=0.19$), but not for the early ($p=0.46$, $r=0.06$) and late ($p=0.79$, $r=0.02$) participants. For the corrugator muscle we found a significant correlation for early participants ($p=0.05$, $r=-0.17$), but not for congenital ($p=0.53$, $r=0.05$) or late ($p=0.15$, $r=-0.13$) groups.
Importantly, all significant EMG correlations were congruent with the evaluated smiliness of the sound.

Finally, to separate the respective contribution of both sound manipulation and participants' ratings in the EMG measures, we performed a GLMM analysis using the EMG data from the 'rating task' (EMG data inside the significant cluster observed at the group level) for each muscle. We used two predictors: sound transformation (2 levels: smile vs unsmile effect), and participants' continuous ratings (numeric factor ranging from 0 to 20); Participant number and sound token were used as random factors.
For the zygomatic muscle, there was a significant main effect of rating ($\chi_2(1) = 11.286$, $p= 0.0007$). Importantly, adding sound manipulation as a predictor to that model significantly improved model's performance ($\chi_2(1) = 6.23$, $p=0.01$). This shows that the acoustic signature of the smile (manipulated here by the audio algorithm) explains a significant amount of variance in the zygomatic data, even when considering participants' co-occurring and explicit rating. For the corrugator muscle, there was no significant effect for either predictor.

**Group analyses – Detection Task**

To understand how facial reactions vary depending on participants' ratings during the detection task, we grouped EMG time series between rating categories (either pronounced « with » or « without » smiles). We then performed cluster permutation analyses for both zygomatic and corrugator muscles. We found two significant clusters, one for the zygomatic muscle ($p=0.006$, peak=2.3, cluster: [1.7, 2.8], $d=0.6$) and one for the corrugator muscle ($p=0.04$, peak=0.36, cluster: [0.01, 0.5], $d=-0.55$), differentiating response categories. Both significant clusters were congruent with the evaluated smiliness of the sound (i.e., higher zygomatic activity and lower corrugator activity for smile ratings compared to no-smile ratings).

We observed no significant EMG differences between smile and unsmile audio manipulations with cluster permutation tests at the group level.

**Group analyses – Rating and Detection Tasks grouped**

We analysed the difference between smile and unsmile EMG activity across the detection and rating tasks with GLMMs, by averaging the EMG time series in the [0.5s - 3s] time range for each trial, each muscle and both experimental tasks for each participant (Supplemental Data S1E). Then, for each muscle, we fitted a GLMM with sound token and participant number as random factors and tested for main effects of sound manipulation (smile vs unsmile), sex (male vs female) and age (numerical variable).
For the zygomatic muscle, we found a main effect of the sound manipulation ($\chi 2(1) = 4.56$, p=0.03). As predicted, the smile manipulation increased zygomatic activity by 1.14 a.u. (± 0.5 SE, p=0.03) when compared to the unsmile manipulation. For the corrugator muscle, the smile manipulation decreased muscle activity, although not significantly ($\chi 2(1) = 1.4$, p=0.24).

We found no main effects of age or sex for the corrugator muscle (sex: $\chi 2(1) = 2e\text{-}04$, p= 0.98; age: $\chi 2(1) = 2.05$, p= 0.15); No main effect of sex for the zygomatic muscle (sex: $\chi 2(1) = 0.53$, p= 0.46), but a marginally significant effect of age ($\chi 2(1) = 3.15$, p= 0.07): young blind participants had a tendency to show stronger zygomatic activity than older participants. However, we did not find significant interactions between age and sound manipulation or between sex and sound manipulation neither for the zygomatic (age: $\chi 2(1) = 2.01$, p=0.15; sex: $\chi 2(1) = 2.15$, p=0.14) or the corrugator muscles (age: $\chi 2(1) = 0.18$, p=0.67; sex: $\chi 2(1) = 0.14$, p=0.70). In sum, these analyses suggest that, in our data, mimicry reactions in blind participants were not significantly influenced by participants' age or sex.

## EMG data – Individual Statistics

Our study is based on a small-N, large-number-of-trial experimental design, in which we compute individual statistics with a large sample size at the participant level, and treat the participant as the replication unit[S13]. At the participant level, we determined sample size using an hypothetical effect size of d=0.3 (where our previous study of sighted controls found d=0.5[S5]) and power = 0.85, yielding 80 matched stimuli (N=240 total trials per participant).

In order to study individual differences in EMG responses to smiliness within the group, we performed cluster permutation tests[S12] for both muscles, and for each participant, by considering all 240 trials (rating and detection tasks grouped). The results for the Zygomatic muscle are presented in Figure S1, the results for the corrugator muscle are presented in Figure S2. All significant clusters are presented in Supplemental Data S1F.
For the zygomatic muscle, EMG activity differences between smile and unsmile stimuli were significant at the individual level for 4 participants, all of which were congenitally blind (4/5, 80%). For the corrugator muscle, EMG differences between smile and unsmile conditions were significant at the individual level in 6 clusters across 4 participants (congenital:1; early:2; late:1). Importantly, across participants and muscles all significant clusters followed the predicted effect direction (congruent muscle reactions: unsmile < smile for the zygomatic muscle; unsmile > smile for the corrugator muscle).

## Comparison with a sighted control group

**Methods:**
In order to compare the facial reactions of blind participants to those usually observed in sighted individuals, we used the data from our previous study investigating facial reactions to auditory smiles in a group of neurotypical adults[S5]. In that study, N=35 participants performed virtually the same experiment as the blind participants of the present study. The task question was the same as the one in the rating task ("to what extent [was] this sentence pronounced with a smile?"), the algorithm for manipulating stimuli was the same, and the EMG sensors, apparatus and pre-processing were the same.

Sighted participants and controls differed on two aspects. First, sighted participants rated stimuli using a mouse and a visual rating scale, whereas blind participants rated stimuli by providing digital inputs with a computer keyboard. To control for this difference, in the following we compare only measures of effect sizes between groups, and not explicit ratings. Second, sighted participants differed in their number of trials with blind participants. Sighted participants performed 60 trials, whereas blind participants performed 240 trials. To create two comparable groups, we randomly generated a set of N=250 simulated "meta-participants" from the data from the 2018 study. To do this, for each participant, we randomly chose 80 matched trials (smile, neutral and unsmile) from the dataset of all sighted individuals. This way, meta-participants were matched in their number of trials with blind participants.

**Smile decoding Accuracy – sighted vs blind**
In order to compare smile decoding accuracy of sighted and blind groups, we computed Cohen's d for each sighted meta-participant and blind participant. To do this, we used the data from the rating-task (120 trials) and paired ratings (smile vs unsmile), normalized by the neutral (non-manipulated) token. Data is presented in Supplemental Data S1G.
The mean effect size was M=0.73 (std : 0.30) for controls, and M=0.54 (std : 0.34) for blind participants. That difference was not statistically significant: Welsh's unequal variance t-test(13.0)= 1.95, p=0.07 (Supplemental Data S1G).

**EMG Data Analysis – sighted vs blind**
In order to compare facial reactions between sighted meta-participants and blind participants, we computed individual statistics at the individual level using cluster permutations tests between smile and unsmile time-series for both zygomatic and corrugator muscles. We then computed Cohen's d measure of effect size for each muscle and for each significant and congruent cluster using the distributions of mean EMG activity inside the significant clusters. For this analysis, we included all clusters where $p < 0.10$, instead of 0.05, in order to have more sensibility to potential variations between groups. Data is presented in Supplemental Data S1G.

We first investigated if blind participants exhibit more individually-significant clusters of mimicry than controls. To do so, we counted the number of participants having at least one cluster differentiating zygomatic/corrugator activity between smile and unsmile trials.
In the sighted meta-participant group, we found a total of 232 clusters that significantly differed between smile and unsmile conditions (133 for zygomatic; 99 for corrugator). Of these clusters 81.4% were congruent with the acoustic manipulation (190 congruent). 35% of meta-participants

had at least one significant and congruent effect at the individual level for the zygomatic muscle (88 participants); 23% for the corrugator muscle (57 clusters).

For blind participants, we found a total of 15 clusters (10 presented in Supplemental Data S1F, and 5 where alpha was between [0.05; 0.10]) where facial reactions congruently differed between smile and unsmile conditions (we found no incongruent clusters in the blind data set). Such clusters were distributed among 5 participants for the zygomatic muscle and 5 participants for the corrugator muscle.
We used these distributions to test whether blind participants exhibit more mimicry than sighted controls. Blind and sighted distributions did not differ statistically neither for the corrugator muscle ($\chi^2$=1.68, p=0.29) nor for the zygomatic muscle ($\chi^2$=0.06, p=1)—$\chi^2$ test statistic and p-value computed with Monte Carlo simulations based on 10000 replications.

We then investigated whether, within clusters of significant and congruent mimicry, blinds exhibit more intense effects than controls. To do so, we measured Cohen's d effect size using the difference in EMG activity between smile and unsmile trials for each significant cluster, muscle and participant, for both sighted and blind groups.
For the corrugator muscle, the mean effect size was M=-0.40 (std: 0.05) for controls, and M=-0.42 (std: 0.09) for blind participants. That difference was not statistically significant: Welsh t(10)=0.83, p=0.42 (Supplemental Data S1G). For the zygomatic muscle, the mean effect size was M=0.39 (std: 0.05) for controls, and M=0.37 (std: 0.03) for blind participants, again, the difference was not statistically significant: Welsh t(4.76)=1.23, p=0.27 (Supplemental Data S1G).

In sum our data suggests that (1) blind participants do not exhibit more individually significant clusters than controls and (2) the intensity of mimicry effects in blind participants is not stronger or weaker than in controls.

Note, however, that the unbalanced sample sizes (N=14 vs N=250) give the statistical tests comparing blind and sighted participants sensitivity only to large effects (power 0.85 for effects d> 0.7). Our design therefore cannot rule out true differences of ratings or EMG reactions between blinds and sighted participants, but we can say that, if they exist, such differences are small/medium (d<0.7). This further reinforces our claim that mimicry is preserved in blind participants, demonstrating that visual experience is not a necessary condition for mimicry to develop.

## Author Contributions

Experimental design, P. A., J.J.A., C.B.; Data collection, P.A., C.B.; Data analysis, P.A, J.J.A.; Writing – Original Draft, P.A. and JJ.A.; Writing – Review & Editing, P.A., J-J.A., C.B.

## Supplemental References


S1. Wan, C.Y., Wood, A.G., Reutens, D.C., and Wilson, S.J. (2010). Early but not late-blindness leads to enhanced auditory perception. Neuropsychologia *48*, 344–348.

S2. Zafeiriou, D.I., Ververi, A., and Vargiami, E. (2007). Childhood autism and associated comorbidities. Brain Dev. *29*, 257–272.

S3. Ponsot, E., Arias, P., and Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. J. Acoust. Soc. Am. *143*, EL19--EL24.

S4. Arias, P., Soladie, C., Bouafif, O., Robel, A., Seguier, R., and Aucouturier, J.-J. (2018). Realistic transformation of facial and vocal smiles in real-time audiovisual streams. IEEE Trans. Affect. Comput.
S5. Arias, P., Belin, P., and Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. Curr. Biol. *28*, R782--R783.
S6. Pampalk, E. (2004). A Matlab Toolbox to Compute Similarity from Audio. In Proceedings of the ISMIR International Conference on Music Information Retrieval, Barcelona, Spain.
S7. Boersma, P.P.G., and others (2002). Praat, a system for doing phonetics by computer. Glot Int. *5*.
S8. Boxtel, A. Van (2010). Facial EMG as a tool for inferring affective states. Proc. Meas. Behav. *2010*, 104–108.
S9. Peirce, J.W. (2007). PsychoPy - Pyschophysics software in Python. J. Neurosci. Methods *162*, 8–13.
S10. Künecke, J., Hildebrandt, A., Recio, G., Sommer, W., and Wilhelm, O. (2014). Facial EMG responses to emotional expressions are related to emotion perception ability. PLoS One *9*, e84053.
S11. Gelman, A., and Hill, J. (2007). Data analysis using regression and multilevelhierarchical models (Cambridge University Press New York, NY, USA).
S12. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods *164*, 177–190.
S13. Smith, P.L., and Little, D.R. (2018). Small is beautiful: In defense of the small-N design. Psychon. Bull. Rev.

## Supplemental Figures

*Figure S1: Zygomatic time series for all participants during the listening of the stimuli; red bars indicate clusters where smile and unsmile time series significantly differ ($p<0.05$, cluster permutation tests); Note that all significant clusters are congruent with the sound manipulation (smile > unsmile). Shaded areas represent SEM.*

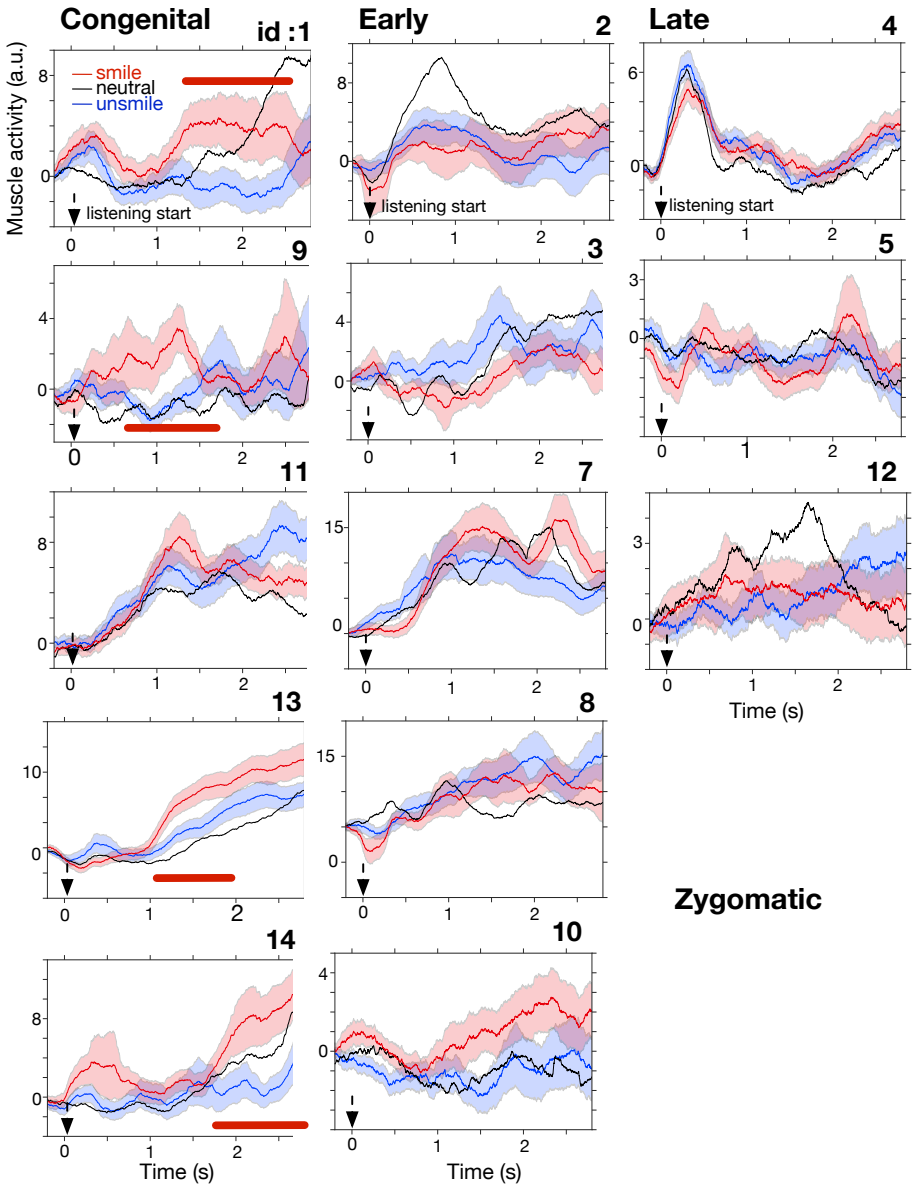
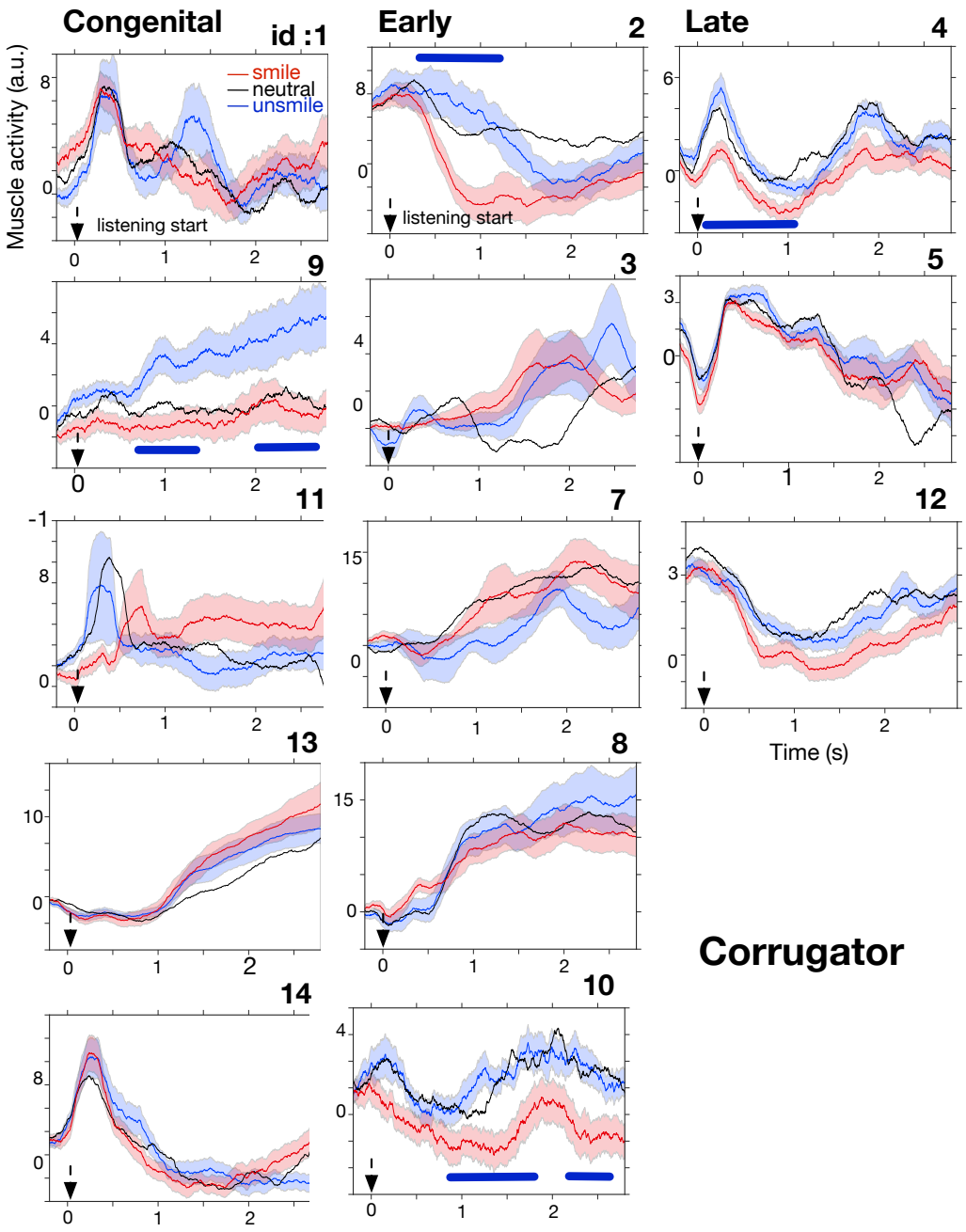



*Figure S2: Corrugator time series for all participants during the listening of the stimuli; blue bars indicate clusters where smile and unsmile time series significantly differ ($p<0.05$, cluster permutation tests); Note that all significant clusters are congruent with the sound manipulation (smile < unsmile). Shaded areas represent SEM.*


Congenital
Early
Late
id :1
2
4
9
3
5
11
7
12
13
8
14
10
smile
neutral
unsmile
listening start
Muscle activity (a.u.)
Time (s)
Corrugator

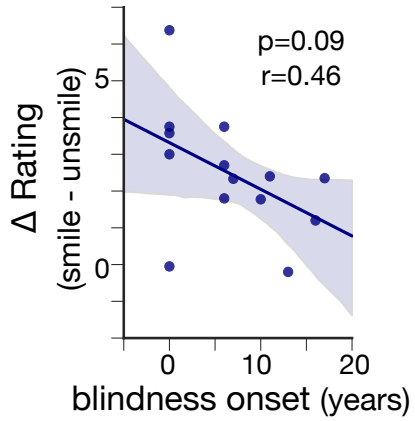
# Supplemental DATA S1

**Complementary information about participants, individual statistics, and group comparisons.**

A) Etiology of blindness in tested participants. ICD-10 is the 10th revision of the International medical Statistical Classification of Diseases and Related Health Problems (ICD)

| Category | Part. id | visual capacities | Etiology | Blindness Onset (in years) | IC10 |
|---|---|---|---|---|---|
| **Congenital** | 1 | light perception | retinal dystrophy | 0 | H35.5 |
| | 9 | light perception | retinal dystrophy | 0 | H35.5 |
| | 11 | 0 | microphtalmus (eye globe abnormality) | 0 | Q11.2 |
| | 13 | light perception | retinal dystrophy | 0 | H35.5 |
| | 14 | 0 | ocular tumor | 0 | C69.2 |
| **Early** | 2 | light perception | retinal disease | 6 | H35.9 |
| | 3 | 0 | optic nerve hypoplasia | 6 | H47.03 |
| | 6 | light perception | congenital cataract | 7 | Q12 |
| | 7 | vague light perception | eye globe abnormality | 6 | Q11.2 |
| | 8 | 0 | retinopathy of prematurity | 10 | H35.1 |
| | 10 | 0 | congenital cataract | 11 | Q12.0 |
| **Late** | 4 | Light perception | Congenital glaucoma | 17 | Q15 |
| | 5 | vague mouvement perception & light perception | Stargardt disease | 13 | H355 |
| | 12 | light Perception | Retinitis pigmentosa | 16 | H35.52 |

B) Correlation between participants' blindness onset and their smile discrimination accuracy; p=0.09; r=0.46 (Pearson correlation coefficient)

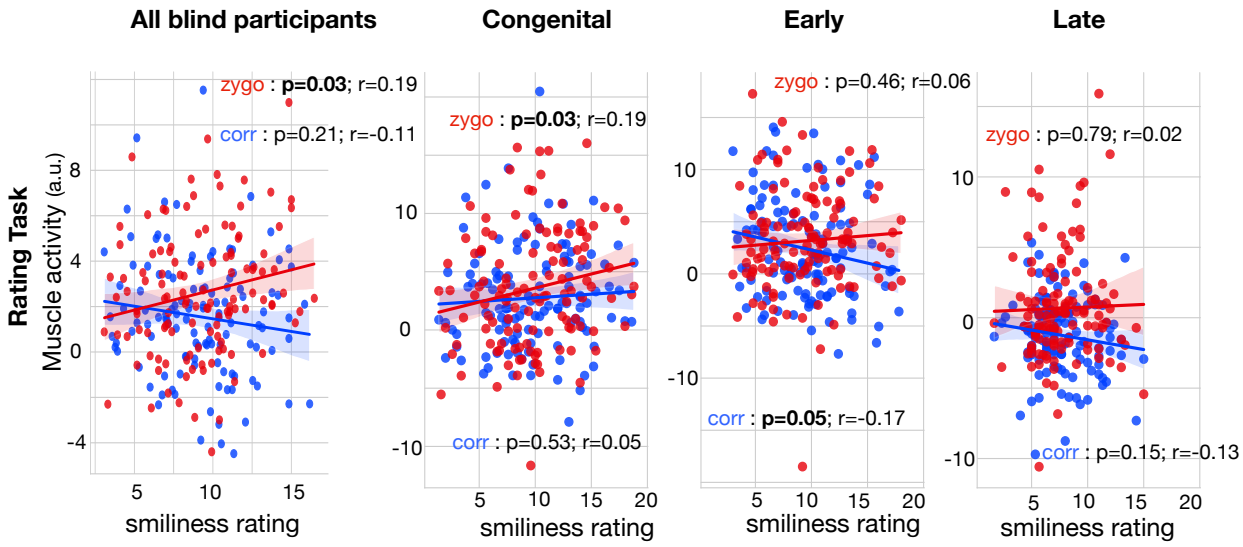



C) GLMM Statistics at the individual level for both the rating and the detection tasks; significant (sign.) codes: .: p< 0.1; * : p<0.05; ** : p< 0.01 ; ***: p<0.001; Cohen's d was computed by comparing paired rating distributions of smile and unsmile trials; Detection accuracy was computed as the percentage of correct responses in the detection task.

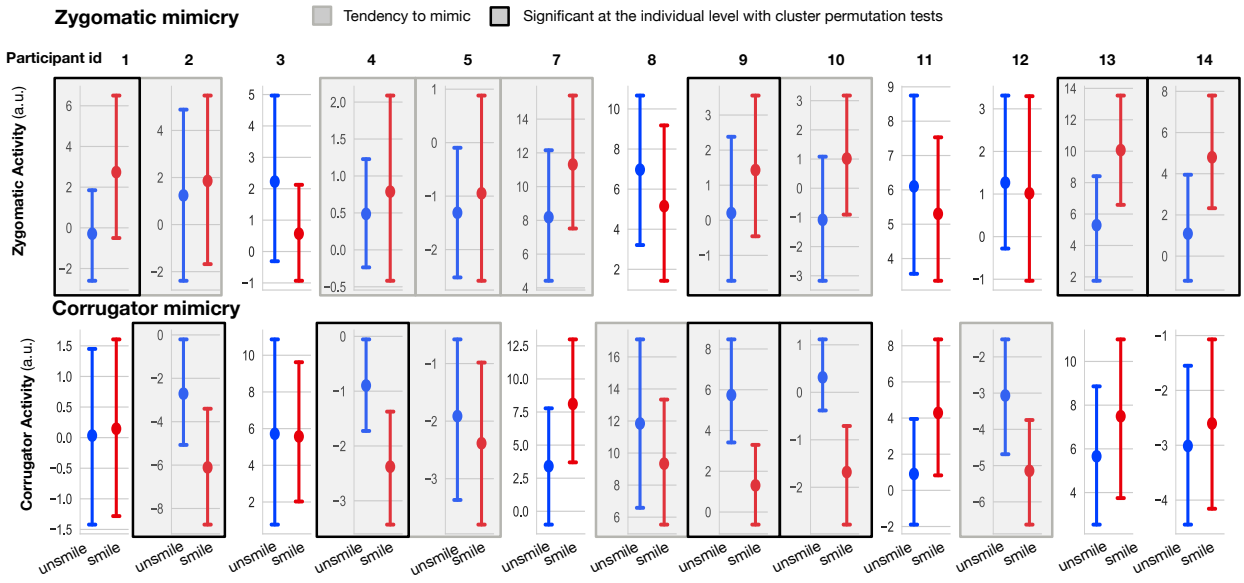
| Category | Part. id | Rating task | Sign. code | Cohen d | Detection task | Sign. code | Accuracy (%) |
|---|---|---|---|---|---|---|---|
| **Congenital** | 1 | $\chi2(5)$ =22.8, p=1e-5 | *** | 0.90 | $\chi2(4)$ =0.4, p=0.8 | | 51 |
| | 9 | $\chi2(5)$ =15.1, p=0.0005 | *** | 0.60 | $\chi2(4)$ =2.5, p=0.28 | | 55 |
| | 11 | $\chi2(5)$ =17.8, p=0.0001 | *** | 0.75 | $\chi2(4)$ =9.3, p=0.009 | ** | 60 |
| | 13 | $\chi2(5)$ =34.4, p=3.3e-8 | *** | 1.1 | $\chi2(4)$ =7.0, p=0.03 | * | 63 |
| | 14 | $\chi2(5)$ =0.9, p=0.6 | | -0.01 | $\chi2(4)$ =0.8, p=0.65 | | 54 |
| **Early** | 2 | $\chi2$ (5) =10.6, p=0.004 | ** | 0.38 | $\chi2(4)$ =18.8, p=8e-5 | *** | 68 |
| | 3 | $\chi2$ (5) =21.0, p=2.6e-5 | *** | 0.69 | $\chi2(4)$ =12.8, p=0.001 | ** | 62 |
| | 6 | $\chi2$ (5) =10.8, p=0.004 | ** | 0.43 | $\chi2(4)$ =11.4, p=0.003 | ** | 63 |
| | 7 | $\chi2$ (5) =7.8, p=0.02 | * | 0.42 | $\chi2(4)$ =4.9, p=0.08 | . | 55 |
| | 8 | $\chi2$ (5) =9.5, p=0.008 | ** | 0.49 | $\chi2(4)$ =11.03, p=0.004 | ** | 66 |
| | 10 | $\chi2$ (5) =24.2, p=5.4e-6 | *** | 0.86 | $\chi2(4)$ =18.2, p=0.0001 | *** | 67 |
| **Late** | 4 | $\chi2$ (5) =16.0, p=0.0003 | *** | 0.65 | $\chi2(4)$ =16.2, p=0.0003 | *** | 60 |
| | 5 | $\chi2$ (5) =0.2, p=0.8 | | -0.08 | $\chi2(4)$ =5.0, p=0.08 | . | 56 |

| | | | | |
|---|---|---|---|---|
| 12 | $\chi 2$ (5) =4.9, p=0.09 | . | 0.34 | $\chi 2$(4) =3.8, p=0.14 | 59 |

D) Rating Task: Correlations between smiliness ratings and EMG activity for both corrugator (blue) and zygomatic (red) muscles for all participants, as well as congenital, early and late participants; r: Pearson correlation coefficient.



E) Rating and Detection Task grouped: Mean zygomatic (top) and corrugator (bottom) activity for each blind participant. Participants shaded in grey show muscle activity that is congruent with the audio manipulation, i.e. high zygomatic for the smile (red) condition compared to the unsmile (blue) condition, or lower corrugator activity for the smile (red) condition compared to the unsmile (blue) condition). Participants showing a significant effect at the individual level with cluster permutation tests are additionally bordered in black (see Individual Statistics section in SI ). Error Bars are 95% Confidence Intervals.




F) Significant cluster permutation tests at the individual level. Cluster: time interval in seconds where 'smile' time series are significantly different from unsmile time series; peak: time tag of the maximum difference between conditions inside the significant time cluster; Direction: whether the direction of the change between smile and unsmiled time series is congruent with the audio transformation.

| Category | Muscle | Participant id | Cluster | p-value | Peak | Direction | Cohen-d |
|---|---|---|---|---|---|---|---|
| **Congenital** | Zygomatic | 1 | 1.2-2.4 | 0.02 | 2.4 | Congruent | 0.33 |
| | | 9 | 0.8-1.4 | 0.04 | 1.2 | Congruent | 0.37 |
| | | 13 | 1.1-1.8 | 0.05 | 1.6 | Congruent | 0.37 |
| | | 14 | 2.1-2.8 | 0.03 | 2.5 | Congruent | 0.42 |
| | Corrugator | 9 | 0.9-1.9 | 0.003 | 1.4 | Congruent | -0.44 |
| | | 9 | 2.0-2.8 | 0.007 | 2.4 | Congruent | -0.36 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Early** | Corrugator | 2 | 0.5-1.3 | 0.02 | 1.0 | Congruent | -0.50 |
| | | 10 | 0.9-2.0 | 0.002 | 1.5 | Congruent | -0.49 |
| | | 10 | 2.0-2.8 | 0.007 | 2.4 | Congruent | -0.43 |
| **Late** | Corrugator | 4 | 0.1-1.1 | 0.0006 | 0.2 | Congruent | -0.56 |

G) Blind vs sighted meta-participants. (G.a) Difference in detection accuracy between sighted meta-participants and blind participants. Each point represents the Cohen's d for an individual participant, computed using smile and unsmile rating distributions (rating task). n.s : not statistically significant Welsh's unequal variance t-test t(13.0) =1.95, p=0.07. (G.b) Effect sizes of significant and congruent clusters of zygomatic (left) and corrugator (right) activity for both sighted and blind participants. Effect size was computed as the Cohen's d of the mean EMG data inside significant clusters (independently detected with cluster permutation tests). n.s : not statistically significant Welsh's unequal variance t-test t(4.76)=1.23, p=0.27 for the zygomatic muscle and t(10)=0.83, p=0.42 for the corrugator muscle.