124 (2025) 202599

Contents lists available at ScienceDirect

Research in Autism

journal homepage: www.elsevier.com/locate/rasd

Intact representation of vocal smile in autism: A reverse correlation approach

Annabelle Merchie [a,b,c], Zoé Ranty [b,c], Aynaz Adl Zarrabi [d], Frédérique Bonnet-Brilhault [b,c,e], Emmanuelle Houy-Durand [b,c,e], Jean-Julien Aucouturier [d,f], Marie Gomot [b,c,*,1]

[a] ACTE at ULB Neuroscience Institute, Université Libre de Bruxelles, 50 Avenue F.D. Roosevelt, Brussels, 1050, Belgium
[b] Université de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraiN U1253, Tours 37032, France
[c] Excellence Center for Autism and Neurodevelopmental Disorders, EXAC·T, CHRU de Tours, Tours, France
[d] Université de Franche-Comté, SUPMICROTECH, CNRS, institut FEMTO-ST, Besançon, France
[e] Centre Ressource Autisme région Centre-Val de Loire, CHRU de Tours, Tours, France
[f] STMS Lab IRCAM, CNRS, Sorbonne Université, Paris, France

ARTICLE INFO

Keywords:
Autism
Prosody
Vocal smile
Reverse correlation
Internal noise

ABSTRACT

Atypical emotional prosody production and perception have been reported in autism. However, it is unclear whether these particularities are associated with unusual mental representations of vocal emotions. The objective of the current study was to explore the mental representation of vocal smile in autistic adults. Twenty-nine autistic (ASD) and 29 neurotypical (NT) adults performed an auditory reverse correlation task, that affords the opportunity to extract acoustic features of mental representation and their variability. Most ASD participants (17) based their representation of vocal smile on similar acoustic features as NT participants and no difference in the level of internal noise was observed. However, comparisons between groups revealed a more typical representation in NT than in ASD. Subsequent cluster analysis revealed that the difference of typicality was explained by a small subset of ASD participants displaying different representations. A correlation analysis also revealed that the typicality was positively correlated with the empathetic level within both groups. While most autistic adults have a preserved mental representation of vocal smiles, a subset shows less robust and typical representations, which is linked to lower levels of empathy. This study highlights that the perception of vocal smiles in autism is more nuanced than previously reported, with empathy playing a substantial role in shaping these mental representations.

1. Introduction

Prosodic atypicalities are considered as a central feature of the communication profile in autism (ASD). While atypicalities in the perception and production of emotional prosody have been widely studied in children (Hubbard & Trauner, 2007; Le Sourn-Bissaoui et al., 2013; McCann & Peppé, 2003; Peppé et al., 2007; Taylor et al., 2015) and adults with autism (Hubbard et al., 2017) since the

* Corresponding author at: Université de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraiN U1253, Tours 37032, France.
E-mail address: gomot@univ-tours.fr (M. Gomot).
[1] ORCID: 0000–0003-1747–5806

https://doi.org/10.1016/j.reia.2025.202599
Received 14 August 2024; Received in revised form 11 March 2025; Accepted 3 April 2025
Available online 13 April 2025
3050-6565/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

original descriptions of Kanner and Sukhareva (Andronikof & Fontan, 2016; Kanner, 1943), the evidence that prosodic processing is inherently impaired in autism remains mixed (Zhang et al., 2022). In particular, it is unclear whether the difficulties processing emotional prosody that are associated with autism result from abnormal sensory representations of vocal emotions per se, or from confounds linked to challenges in social interactions (Scheerer et al., 2020) and other higher-level cognitive processes (Day et al., 2023). Beyond studying basic sensory levels, no study has yet tested whether the representation of vocal emotion in ASD was based on relevant acoustic cues, and how stable they were.

An early hallmark of autism is also a lack of response to (Beall et al., 2008) and production of (Dawson et al., 1990) social smile, one of the most emblematic expressions of the human emotional repertoire. Understanding and evaluating mental representation of vocal smile in autism may provide key insights into the nature of prosodic atypicalities in autism.

Behavioral studies in autism have shown difficulties in recognition of others' prosodic intention (see Leung et al., 2022 for a review). However, pitch discrimination in autism has been described as better than in neurotypical individuals for non-vocal sounds (Bonnel et al., 2003, 2010; Heaton et al., 2008) with a larger proportion of absolute pitch in the autistic population (Miller, 1999; Rimland & Fein, 1988). Nevertheless, these improved auditory skills in autism are not sufficient for efficient recognition of vocal emotions, suggesting that atypical prosodic processing does not arise from sensory processing difficulties, but rather from subsequent integrative processes enabling the attribution of meaning to specific voice patterns.

Perceptive representation of several stimulus categories have been extracted from physical features through the reverse correlation technique (i.e., a data driven exploratory method to extract features used in the classification of a stimulus), historically in the visual modality (Mangini and Biederman, 2004), but has recently gained in influence in the auditory modality (Adl Zarrabi et al., 2024; Goupil et al., 2021; Merchie et al., 2024; Ponsot, Arias et al., 2018; Ponsot, Burred et al., 2018; Wang et al., 2022). Smile could be identified from vocal cues (Aubergé & Cathiard, 2003; Tartter, 1980; Tartter & Braun, 1994) and thanks to reverse correlation, a robust acoustic representation of vocal smile was revealed, which even led to motor resonance in neurotypical adults (Arias et al., 2018; Merchie et al., 2024). Vocal smile internal representation is characterized by an upward shift in frequency of formants F1 and F2, and an increase in energy of F2, F3 and F4 compared to neutral voice (Ponsot, Arias et al., 2018), reflecting the modulation of vocal tract caused by the bilateral contraction of the Zygomaticus major muscles during smile (Wood et al., 2016). Few studies have used reverse correlation techniques to characterize mental representations in autism, either in the visual or auditory modality. In a visual reverse correlation "Bubbles" task, results suggested that autistic children did not use the same information than neurotypical during identity judgment (Ewing et al., 2018). However, another study demonstrated that to identify positive emotion, contrary to eye tracking studies (Black et al., 2017), ASD children correctly used the information from the eyes region (Song et al., 2012). Only one study used an auditory reverse correlation paradigm in autistic children and observed similar representations of non-emotional pitch contours in a tonal language (Mandarin), and musical melody contours (Wang et al., 2022). As yet, no study has investigated the evaluation of acoustic cues used in the classification and representation of vocal emotion in autism.

From another perspective, difficulties in processing vocal sounds in autism highlight the importance of considering the stability of the associated representations as key factor. It is well known that human sensory processing is inherently noisy (Neri, 2010), with typical levels of additive "internal noise" mirroring or even slightly exceeding the variance found in the natural environment, which has been suggested to decrease sensory thresholds and facilitate generalization and adaptation to changes in the surroundings (Faisal et al., 2008). While both neural and behavioral within-subject variability have been proposed to contribute to sensory symptoms in ASD (Magnuson et al., 2020; Ward, 2019), whether this variability also impacts internal representations remains to be explored. In reverse correlation studies, the internal noise referred partly to "intrinsic noise in the decision making" and could be estimated in reverse correlation task through the double pass-paradigm as the within-subject response variability (i.e., repetition of 2 blocks of identical stimuli to estimate response variability) (Ponsot, Arias et al., 2018). Few studies used the double-pass methodology to estimate response variability in autism, but they reported an increased level during a visual task (Park et al., 2017; Vilidaite et al., 2017) or no difference with an audio reverse correlation protocol compared to neurotypicals (Wang et al., 2022).

The main objective of the study was to determine whether perceptive representation of the vocal smile in autistic adults was based on similar relevant acoustic cues as those of neurotypicals. The perceptive representations of vocal smile were assessed in a group of autistic adults with an adapted reverse correlation paradigm and compared to the robust neurotypical representation (Merchie et al., 2024; Ponsot, Arias et al., 2018), through the calculation of the representation typicality (i.e., standardized distance between representations) taking as a reference Ponsot et al.'s representation (2018). This enabled us to determine whether autistic individuals rely on the same acoustic cues as those of neurotypicals to classify a voice as smiling, by extracting for each participant an individual representation to be compared to the reference. This paradigm also allowed to estimate the robustness of the representation by extracting within-subject response variability, referred to as internal noise, thanks to double-pass methodology. Analysis of the combination of mental representation typicality and robustness indices should give the opportunity to identify different profiles in autistic adults and to link these profiles to clinical information, therefore reflecting more accurately ASD heterogeneity.

If ASD participants display abnormal representations (i.e. higher representation atypicality than controls), this would suggest that impairments in emotional prosody recognition result from attending to irrelevant acoustic cues, possibly badly generalized from previous exposures. If, on the contrary, ASD have typical representations but abnormal levels of internal noise (i.e. within-subject response variability statistically larger than controls), this would suggest that impairments in vocal emotion recognition mainly relate to a systemic unability to use them consistently, because of higher-level difficulties linked to decision-making and executive functions. If, finally, ASD individuals do not differ from controls on either representation and noise, this would suggest that impairments in prosody recognition do not result from sensory/perceptual mechanisms, but possibly in difficulties engaging these representations in pragmatic demands such as real-world social interaction which are absent from the current experiment.

2. Material and methods

2.1. Population

Thirty autistic and 29 neurotypical participants participated in the current study. Autistic adults were recruited through the Autism Resource Center of Centre Val de Loire and diagnosed by a trained multi-disciplinary clinical team according to DSM-5 criteria. Diagnosis was confirmed with ADOS-2 and ADI-R (Lord et al., 1994, 2000). The ADOS scores for eight autistic participants were slightly below the threshold of 6; however, their diagnosis was confirmed by the clinical team. Additionally, the participants with the lowest ADOS scores had an ADI-R score that exceeded the clinical threshold. Verbal and non-verbal efficiencies were estimated through the use of selected subtests of the WAIS-IV (Wechsler, 2012) for neurotypicals, while the entire scale was proposed to autistic participants. The Autism Quotient 50-questions (AQ) and the Empathy Quotient 40-questions (EQ) were proposed to all the participants (Baron-Cohen et al., 2001; Baron-Cohen & Wheelwright, 2004). One participant was discarded in the autistic group because of a doubt about diagnosis and another participant who was not able to complete the task. The group of neurotypical adults reported an absence of any developmental difficulties in language and sensorimotor acquisition. In both groups, no disease of central nervous system, infectious or metabolic diseases, epilepsy, or an abnormal audition was reported. Participant's audition was checked for both ear with a short subjective audiometry test. Every participant (or their legal guardian in case of adult under guardianship) signed an informed consent form, and the protocol received approval from Ethic Committee (PROSCEA 2017/23; ID RCB: 2017-A00756–47). A description of the two groups is presented in Table 1.

2.2. Sound modulations

The same method and stimuli as in Ponsot et al. (2018) have been used to evaluate mental representation of vocal smile in both groups. In brief, the French phoneme /a/ uttered by a male speaker with a constant pitch was recorded. To obtain a constant spectral energy a 500 ms stationary part of the sound was selected. To generate spectral variants, the original sounds was modulated using a random frequency equalizer. This equalizer consisted of 25 log-spaced frequency points (between 100 and 10,000 Hz) with linearly interpolated gain values (in dB). Gains were drawn from Gaussian distributions (SD = 5 dB) and clipped at ± 2.5 SD with the Cleese python toolbox (Burred et al., 2019).

2.3. Procedure

The current experiment took place in a larger session (~1h30) of testing for most of the participants. The sequence of stimulation for the reverse correlation task was composed by 200 trials for autistic (~15 min) and 300 trials for neurotypical (~20 min) divided in blocks of 50 trials.

To assess the optimal number of trials for the reverse correlation task, an initial study was conducted with neurotypical adults using 250 trials, supplemented by an additional 50 trials at the end of the sequence for double-pass variability estimation (Merchie et al., 2024). In this study, the perceptual representation of the vocal smile derived from 250 trials was compared with the norm established by Ponsot et al. (2018), based on 500 trials. The two representations showed no significant differences. Subsequently, a convergence analysis was performed to estimate the minimum number of trials required to produce a robust representation. This analysis revealed that 150 trials were sufficient for capturing a reliable representation of the vocal smile. Based on these findings, a shorter version of the task with 150 trials for autistic adults was chosen and analyzed only the first 150 trials for neurotypical adults who completed the longer task (250 trials).

Each trial consisted of a pair of randomly-filtered voice to be compared to estimate the most smiling voice between both presented. Participants had to choose the most "smiling" sound of each presented pair by pressing a key. The last two blocks were identical to

Table 1
Characteristics of Neurotypical (NT) and Autistic (ASD) group (mean ±standard deviation). vIQ: verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from vocabulary and similarities subtests in NT. nvIQ: non-verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from matrix reasoning and cubes subtests in NT. ADOS: score calculated as the sum of the subdimensions communication and social interaction behaviors. The values in brackets correspond to the number of values included in the calculation of the mean and standard deviation due to missing data.

	NT	ASD	Comparison
Sex	12 ♀ 17♂	7 ♀21♂	$\chi^2 = 1.06$; p = .30
Age	30.1 ± 12	35.3 ± 13	t(54) = −1.55; p = .13
vIQ	124.2 ± 11 (24)	118.3 ± 20 (25)	t(40) = 1.30; p = .20
nvIQ	106.3 ± 12 (24)	100.7 ± 18 (25)	t(44) = 1.32; p = .19
AQ	16.8 ± 7 (23)	34.9 ± 9 (25)	t(47) = −8.30; p < .0001
EQ	43.2 ± 9 (23)	23.7 ± 13 (25)	t(45) = 6.31; p < .0001
ADOS	/	8.8 ± 5.5	/

evaluate internal noise through the double-pass methodology.

2.4. Reverse correlation analysis

2.4.1. Mental representation

The mental representation of vocal smile was modelled for each participant with the classification image technique (Brinkman et al., 2017). The mean pitch contour of voice classified as "unsmiling" was subtracted from the mean pitch contour of "smiling" voice. Then, the resulting representation was normalized by dividing it by the sum of its absolute value. Each participant's representation was then computed and averaged by group. A comparison between NT and ASD's gain was performed in each modulated frequency. The classification image technique was chosen over a GLM approach (Knoblauch & Maloney, 2008) to more accurately estimate the representation of the vocal smile in the case of observers with high variability (Neri, 2011).

To estimate the deviation of participant representation to the representation of an external control group (Norm) (Ponsot, Arias et al., 2018) the distance between the average group representation and the participant was computed and then normalized in both groups ('representation typicality') (see Eq. (1)). The Norm group was composed according to the description in Ponsot, Arias et al. (2018) by 10 participants (5♀5♂, age range: 18–29 years) that performed the same task as in the current study, but with 600 trials.

$$distance_{participant} = \sum_{f=100}^{10,000} \left| Gain_{f,norm} - Gain_{f,participant} \right| \tag{1}$$

$$typicality_{participant} = 1 - \frac{distance_{participant} - \min(distance)}{\max(distance) - \min(distance)} \tag{2}$$

Eq. 1: (1) Calculation of the distance between the Normative representation and one participant's representation as sum of the absolute difference of representation's gains for each of the 25 modulated frequencies (f) between 100 and 10,000 Hz. (2) Calculation of the typicality for each participant as the normalization of the distance between 0 and 1 within NT and ASD groups.

2.4.2. Internal noise

The same method as in Adl Zarrabi et al. (2024) has been used to evaluate internal noise (expressed in units of the standard deviation of stimulus noise) from response consistency (probability of agreement, p-agree) and response bias (p-int1) between the two repeated blocks using the simulation procedure of Neri (2010). Then for each participant a reverse model was applied to obtain the value of internal noise minimizing the error between the observed and the simulated values for a participant's consistency and bias (Internal Noise). As in Neri (2010) internal noise was estimated in the [0; 5 std] range to avoid unreliable large value.

Internal noise was computed using the open-access python toolbox "palin" (https://github.com/neuro-team-femto/palin). Data from one autistic participant were removed for internal noise estimation because the repeated block was not completed.

2.5. Statistical analysis

Statistical analyses were conducted on RStudio 4.0.4 (R. C. Team, 2021; Rs. Team, 2020) with the packages *ggplot2* (Wickham, 2016), *ggpubr* (Kassambara, 2020), *dplyr* (Wickham et al., 2021), *tidyverse* (Wickham et al., 2019) and *factoextra* (Kassambara & Mundt, 2017).

In order to compare mental representation and noise between groups, Student's tests were performed for representation typicality, p-agree, p-int1 and internal noise. Pearson's correlation between scores in questionnaires (AQ and EQ) or clinical scores (ADOS score, the sum of communication and social interaction) and the different indices of the reverse correlation results were also performed. A global correlation analysis was performed when scores were available for both groups to reflect the effect of the autistic traits and empathy in a continuum, after Fisher's z-tests have been performed to confirm that there was no statistical difference between groups in correlation coefficients. As multiple correlations were calculated the p-value was Bonferroni's corrected and the corrected values were reported.

2.5.1. Cluster analysis

To evaluate the potential different profiles in participants a cluster analysis based on both mental representation (representation typicality) and noise (internal noise) was performed. Variables were scaled for normalization for both groups. A K-means clustering method was chosen (see Manenti et al., 2024 for a detailed description of the methods to determine the best number of clusters). The suitability of clustering for this dataset was assessed using the Hopkins statistic. With a value of 0.72, which exceeds the 0.5 threshold, the dataset was deemed appropriate for clustering. To determine the number of clusters, 30 clustering methods were compared using the R library *NbClust* (Charrad et al., 2014). For our data, the most suitable number of clusters was 3. A qualitative description of the different clusters obtained was presented.

3. Results

3.1. Reverse correlation analysis

3.1.1. Mental representation of vocal smile

A Student's test has been performed on each frequency between groups (ASD vs NT) to evaluate potential differences in mental representation of vocal smile in acoustic description (Fig. 1). As the number of comparisons was quite large (25 frequencies) a Bonferroni correction was applied. No difference between the mental representation in ASD and NT has been revealed.

In both group representation of vocal smile was characterized by an upward switch in frequency of F1 and F2, and by an energy increase of F2 and F4.

3.1.2. Representation typicality

Representation typicality was calculated according to the distance to the representation of the Norm group with 600 trials (Ponsot, Arias et al., 2018). The typicality of the representation appears significantly different between NT and ASD group ($t(52) = 2.20$; $p < .05$, see Fig. 2a), suggesting that mental representation of the NT was closer to these of the external Norm than the one of the ASD. Importantly, the kernels in this study do not reflect an internal auditory smile prototype but rather the spectral manipulations that maximize smile perception. Given the human voice's frequency range (100–10,000 Hz), some kernel frequencies may not correspond to perceptually relevant frequencies or contain significant energy. In theory, inaudible frequencies should not systematically influence responses, but due to the limited number of trials—especially in patient populations—statistical artifacts may arise. To ensure a meaningful interpretation, a focus on formant frequencies was done, which are both theoretically and empirically established as key drivers of vocal smile perception. Nevertheless, participants with autism used relevant frequency modulations to classify a vocal smile, and the typicality differences observed is mainly related to irrelevant frequency modulations (< F1 frequency, 555 Hz).

To evaluate if a relation between clinical scores and the representation typicality existed in the ASD group, a correlation analysis was performed with the ADOS score but did not reveal any significant result ($r = -.03$, $p = .91$). Neither was the correlation coefficient between typicality and AQ in NT ($r = .03$, $p = .87$), nor in ASD ($r = -.09$, $p = .66$) significant. Fisher's z-tests conducted to compare the correlation coefficients between two independent samples did not reveal any statistical difference ($z = .42$, $p = .68$).

The correlation between representation typicality and EQ in NT and in ASD were respectively $r = .31$ ($p = .13$) and $r = .21$ ($p = .32$). The Fisher's test revealed no statistical difference between the two correlation coefficients ($z = .38$, $p = .71$). Considering that there was no significant difference in correlation coefficients between groups, a joint analysis of the correlation between AQ or EQ scores and typicality was conducted on the whole population. There was no correlation with AQ score ($r = -.22$, $p_{corr} = .26$) but a significant correlation between the EQ score and the typicality was identified ($r = .34$, $p_{corr} = .03$, see Fig. 2b). The higher the EQ score, the more similar the acoustic smile representation is to that of the Norm group.

3.1.3. Internal noise

To evaluate differences between NT and ASD internal noise, Student's tests were performed to compare p-agree, p-int1 and the estimated internal noise. These comparisons revealed no difference for p-agree ($t(53) = .66$; $p > .1$), internal noise ($(37) = -1.29$; $p > .1$, see Fig. 3a) and p-int1 was revealed ($t(54) = 1.79$; $p < .1$).

Then, as for the typicality, correlation analysis between internal noise and ADOS, AQ and EQ were performed with the same

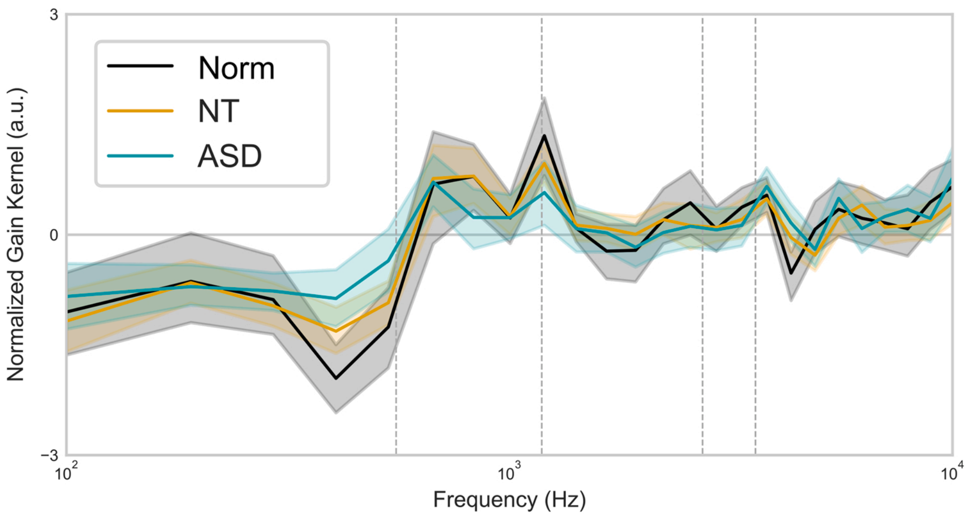

Fig. 1. Mental representation of vocal smile in ASD (blue) and NT (orange). The "Norm" group (black) corresponds to data from Ponsot et al. (2018) with 600 trials. Dashed lines represent formants frequencies.

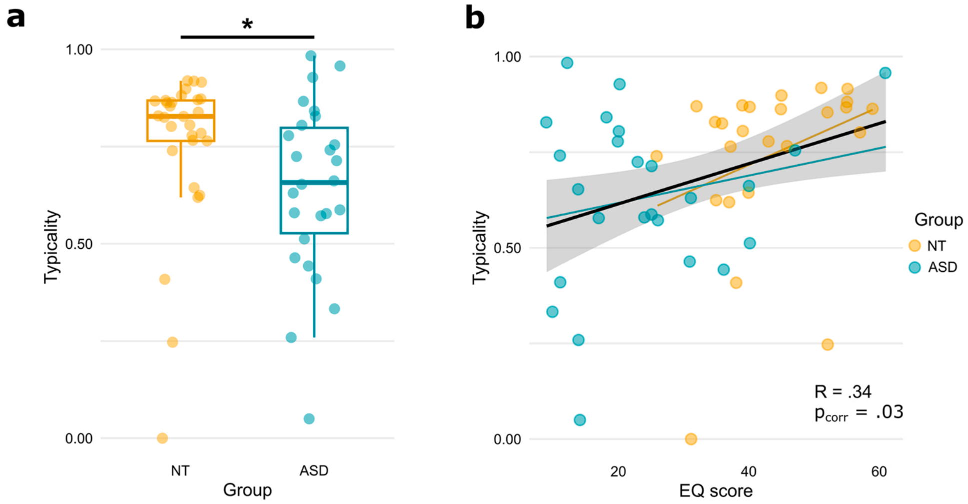

Fig. 2. Typicality of the mental representation of vocal smile. a: typicality of the vocal smile representation to the Norm in NT (orange) and ASD (blue) group * p > .05; b: correlations between typicality and EQ in ASD and NT groups, the fine colored lines correspond to separated group analysis, the bold black line is the correlation within groups.

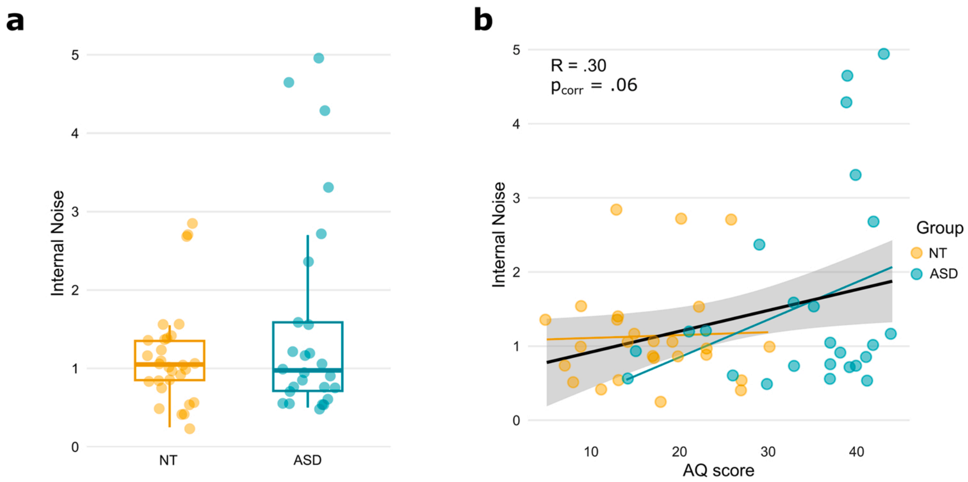

Fig. 3. Internal noise estimation. a: internal noise in NT (orange) and ASD (blue) group; b: correlations between internal noise and AQ in ASD and NT groups, the fine colored lines correspond to separated group analysis, the bold black line is the correlation within groups.

method, and none reached significance. No correlation with ADOS score in ASD group was observed ($r = .04$, $p = .87$).

The correlation coefficient between internal noise and AQ in NT was $r = .04$ ($p = .85$), and in ASD was $r = .32$ ($p = .12$). No statistical difference between the two correlation coefficients, $z = -.97$, $p = .33$ was revealed.

The correlation coefficient between internal noise and EQ in NT was $r = -.43$ ($p = .03$), and in ASD was $r = -.06$ ($p = .77$), but not different statistically ($z = -1.29$, $p = .19$). Considering that there was no significant difference in correlation coefficients between groups, a joint analysis of the correlation between AQ or EQ scores and internal noise was conducted. A tendency to a positive correlation between the internal noise and the AQ was observed ($r = .30$, $p_{corr} = .06$, see Fig. 3b), but there was no correlation with EQ score ($r = -.24$, $p_{corr} = .18$).

3.2. *Cluster analysis*

3.2.1. *Clusters definition*

The optimal number of clusters was automatically estimated to be three in our sample with the k-means method with a Hopkins statistic of 0.72 (above the threshold 0.5) with representation typicality and internal noise (IN) as factors. The profiles were qualified as: *Typical representation – Low IN*, *Atypical representation – Low IN* and *High IN* and are described in Table 2. Clusters are presented with an “illustrative” patient for each cluster in Fig. 4.

Table 2
Description of clusters in ASD and NT (mean ±standard deviation). The values in brackets correspond to the number of values included in the calculation of the mean and standard deviation due to missing data. IN: internal noise.

	Typical representation - Low IN		Atypical representation - Low IN		High IN	
Group	**NT**	**ASD**	**NT**	**ASD**	**NT**	**ASD**
n	23	17	3	6	3	4
Age	31.4 ± 13.6	32.3 ± 12.1	25.6 ± 1.4	41.4 ± 13.2	24.7 ± 2.8	39 ± 13.2
vIQ	124.4 ± 11.5 (18)	117.4 ± 18.8 (14)	119.7 ± 12.1	117.7 ± 22	127.3 ± 8.6	120.5 ± 30.6
nvIQ	105.9 ± 13.9 (18)	103.8 ± 16.8 (14)	108.7 ± 4.2	95.8 ± 20	106.7 ± 4.6	94.5 ± 23.9
AQ	16.6 ± 7 (20)	33.5 ± 9.0 (16)	15 ± 6.2	35 ± 10.9	19.7 ± 6.5	40.3 ± 2.1
EQ	45.2 ± 8.5 (19)	25.2 ± 13.8 (16)	40.3 ± 10.7	20.5 ± 10.4	33.3 ± 6.7	22.5 ± 13.5
Internal noise	1 ± 0.4	1 ± 0.5	0.8 ± 0.3	1.1 ± 0.8	2.8 ± 0.1	4.3 ± 0.7
ADOS	/	8 ± 5.2	/	10.7 ± 5.5	/	9.3 ± 7.1
Representation typicality	0.8 ± 0.1	0.8 ± 0.1	0.2 ± 0.2	0.3 ± 0.2	0.7 ± 0.1	0.5 ± 0.1

As NT and ASD were considered together in the cluster analysis, the distribution of group in each cluster was compared and no significant association between group and cluster distribution was observed ($p = .37$; Fisher's Exact Test; FET).

Comparisons of clinical scores for ASD patients were performed between clusters but did not reveal any statistical differences.

3.2.2. Qualitative description of clusters

Considering the relatively small size of the clusters, a qualitative description was performed to evaluate the differences and the clinical implication of these three profiles (Table 2). Twenty-one patients received a diagnosis of Asperger Syndrome (ICD-10 code: F84.5) indicating no general delay or retardation in language or cognitive development, while six participants were diagnosed with childhood autism (F84.0). The cluster analysis showed that patients with a diagnosis of F84.0 were distributed evenly across the 3 profiles.

First, the *Typical representation – Low IN* group includes almost the entire group of neurotypical adults (23/29) and a majority of the ASD (17/27) group suggesting that, as in Ponsot et al. (2018), the internal representation of the vocal smile is stable and robust in most NT and ASD adults. However, the presence of a small number of typical adults with a different vocal smile profile and associated internal noise highlights the heterogeneity that also exists within the general population. No significant difference in clinical scores was observed between clusters, but the size of the two alternative clusters (*Atypical representation – Low IN* and *High IN*) did not allow to perform correct statistical analysis.

The representations of vocal smile in the *Typical representation – Low IN* revealed that the acoustic cues used to identify a vocal smile were very similar regardless the group (Fig. 5a). In contrast, the *Atypical representation – Low IN* representations were noisy and did not base their representation of smile on the same acoustic cue. The autistic participants in this group used the inverse information to classify a sound as smiling in comparison to the Norm group (Fig. 5b). Finally, in the *High IN* cluster, acoustic cues were not that different from the Norm representation but considering the high internal noise in these participants, interpreting their representations remains challenging (Fig. 5c).

The clinical profiles of the autistic participants did not align with the distribution across clusters. Specifically, when examining patients representing different profiles of vocal smile and internal noise levels, the following was observed: the patient with a *Typical representation* (kernel typicality = 0.98) and a *low IN* (IN = 1.6) had an ADOS score of 18, while patient 29, with *high IN* (IN = 4.6), had an ADOS score of 7. Thus, the ADOS scores do not correspond to the distinct profiles identified in this study.

4. Discussion

This study was the first to characterize mental representation of emotional prosody, a particularly challenging feature, in autism using an emotional auditory reverse correlation paradigm. Results revealed that a large majority of autistic participants based their representation of vocal smile on the same acoustic features as neurotypicals do. However, differences in the perceptive representation of vocal smile in a subgroup of autistic adults in comparison to neurotypical, but also a larger internal noise in another subgroup of autistic participants, have been observed.

4.1. Preserved mental representation of vocal smile

The majority of autistic participants (17/27) used the same acoustic cues in their mental representation of vocal smile as neurotypical, with a robust representation (low internal noise). This result suggests that the representation of the vocal smile is shared by most individuals and preserved in most autistic adults. This is in line with studies that reported no difference in the recognition of emotional prosody neurotypical and autistic participants (Brennand et al., 2011; Brooks & Ploog, 2013; Chevallier et al., 2011; Grossman et al., 2010; Paul et al., 2005). However, some other studies showed difficulties in emotional prosody recognition in autism (Hubbard et al., 2017). Cluster analysis revealed a subgroup of autistic participants with an atypical representation of vocal smile but

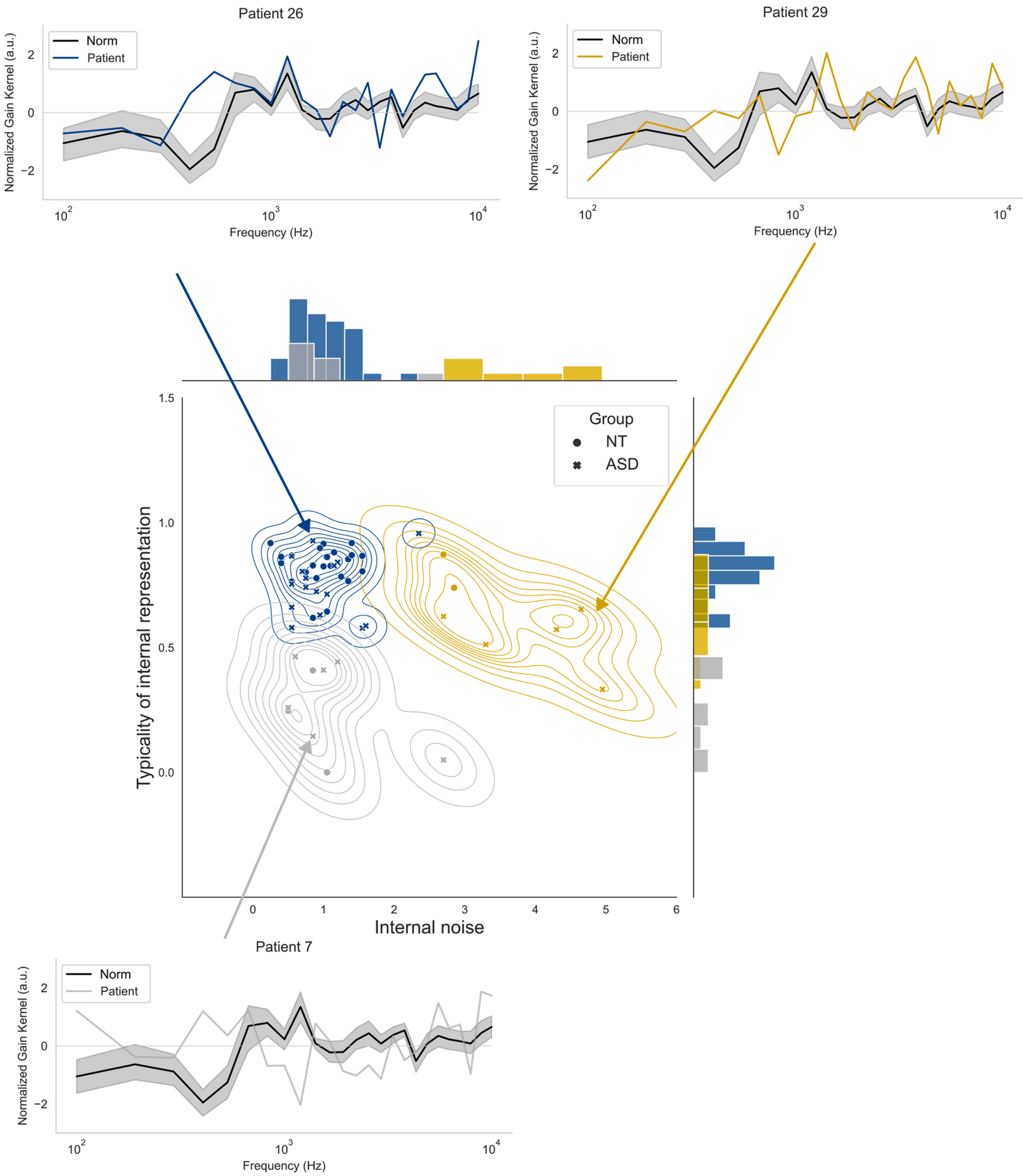

Fig. 4. Cluster analysis according to typicality and internal noise in ASD (cross point) and NT (dot point) and an example of a representative patient of each cluster in the chosen color of the clusters. Typical representation – Low IN (blue); Atypical representation – Low IN (grey) and High IN (yellow). The black curve represents the Norm representation of vocal smile. IN: internal noise.

limited internal noise. This might contribute to the lower typicality of the internal representation of smile in the ASD group as a whole. These findings also suggest that the difficulties in emotion recognition reported in some studies could stem from the heterogeneity of representations in certain autistic participants.

The current study estimates mental representation of vocal smile based on low-level acoustic analysis, and previous research has shown that autistic adults exhibit better skills than neurotypicals in non-vocal acoustic processing (Bonnel et al., 2003, 2010; Heaton et al., 2008). One could hypothesize that studies which report low emotional prosody recognition in autism, used more ecological

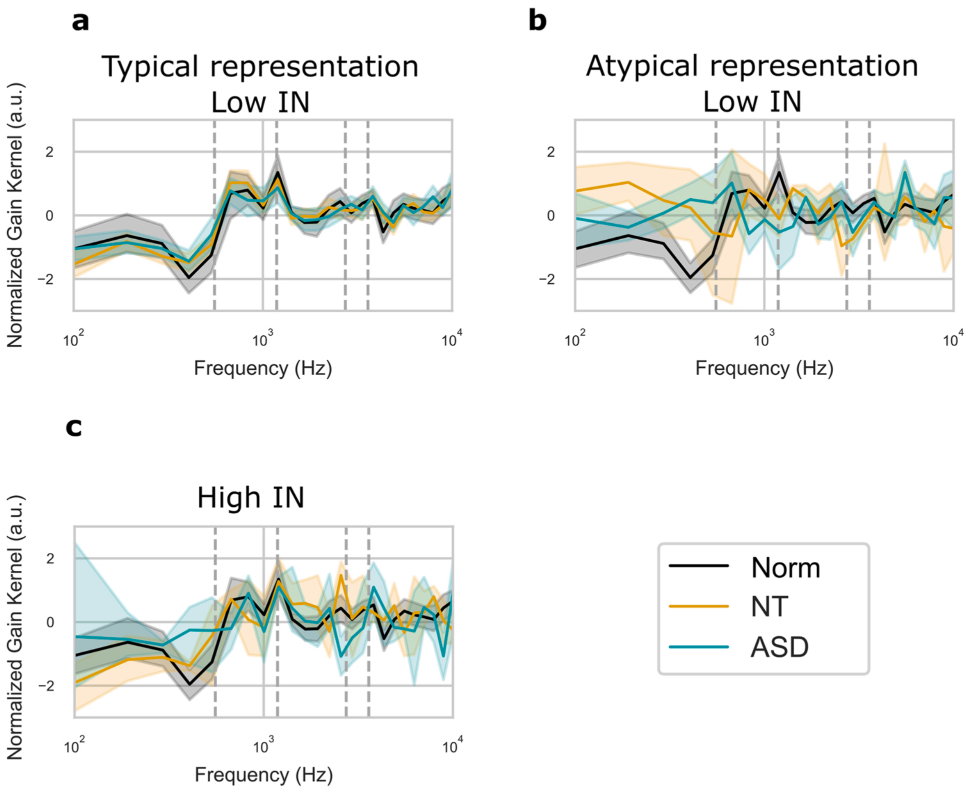

Fig. 5. Mental representation of vocal smile in each cluster for NT (orange) and ASD (blue) in comparison to the Norm representation (black) with 600 trials. IN: internal noise a: Typical representation and low IN cluster; b: Atypical representation and low IN cluster; c: High In cluster.

stimuli and tasks that required higher-level processes than in the present study. The current results demonstrated that the low-level representations of emotional prosody are intact in autism, but one cannot exclude that more spontaneous mechanisms as automatic recognition and judgment could be atypical.

In a previous study, despite reported difficulties in recognizing happiness prosody, a positive association was found between good recognition of happiness prosody and better social adaptation (Wang & Tsao, 2015). While this relationship was not observed between representation typicality and ADOS scores in the autism group in the present study, a positive correlation was found between empathetic abilities and representation typicality. These relationships demonstrated the importance of low-level acoustic processing of emotional prosody in social abilities. It is possible that empathetic abilities, that reflect the ability to understand and share the feelings of other, refine the emotional prosody acoustic processing in neurotypical and autistic adults, or conversely that a typical representation of emotional prosody allows better relations to others and thus develop greater empathy.

Research suggests that autistic individuals may face challenges with audio-visual integration in speech (Hoffmann et al., 2023), which could hinder their ability to associate speech sounds with corresponding articulatory cues (Iarocci & McDonald, 2006; Righi et al., 2018; Kissine et al., 2021). Note however that, contrary to what has been shown in children (De Gelder et al., 1991), recent findings indicate no difference in the McGurk effect between autistic and neurotypical adults suggesting an effect of compensation strategies in adulthood (Jertberg et al., 2024). It remains that these challenges may very well extend to the processing of vocal smile cues, which depend on the integration of auditory and articulatory information (Arias Sarah et al., 2023). This could help explain why some participants rely on different acoustic cues to identify vocal smiles.

4.2. Internal noise

At group level, no difference in internal noise was observed between NT and ASD participants. This lack of difference is compatible with some theories for which noise is not different in autism (Davis & Plaisted-Grant, 2015), but also with the results of a previous reverse correlation study in auditory modality in autistic children in which no difference was shown (Wang et al., 2022).

Nevertheless, cluster analysis revealed a subgroup of participant with a *High IN*. In this profile the mental representation of vocal smile is similar to Ponsot’s results (2018) suggesting that i) unstable noisy representation could lead to both prosody difficulty and hypo- or hyper-sensitivity in autism and ii) the magnitude of internal noise is not related to the quality of the representation of vocal smile. Indeed, for some authors a larger internal (neural) noise might facilitate the detection and discrimination of signal (Davis & Plaisted-Grant, 2015). Nonetheless, in this profile the larger IN (> 3) were observed in four autistic participants only, a result that should be interpreted with caution. The heterogeneous results between participants with autism in terms of internal noise levels also reflect the diversity of participants' clinical profiles.

Finally, in the context of the Bayesian brain, internal noise level and mental representation modify the perception of the environment (Haker et al., 2016). The present study provided an opportunity to estimate an outline of the mental representation to which the percept is compared, in a Bayesian context, and of the noise that would modulate this representation. In autism the Bayesian theories proposed that the sensory input and the prediction (mental representation) differ in weight, which leads to hypo- or hyper-sensitivity (Brock, 2012; Haker et al., 2016; Lawson et al., 2014; Van de Cruys et al., 2014). The level of noise in the signal and internal noise have been proposed to influence the level of prediction of the sensory input according to the context, and thus noisy inputs would lead to hyposensitivity and precise input to hypersensitivity (Van de Cruys et al., 2017).

4.3. Limitations and perspectives

In the present study, the available clinical measures (ADOS, AQ) did not explain the different profiles. The addition of hyper- and/or hypo-sensitivity assessment, as the Dunn sensory profile (Kern et al., 2007) or the Glasgow Sensory Questionnaire (Robertson & Simmons, 2013) for example, would provide complementary information on the internal noise levels observed. Moreover, to draw stronger conclusions on the association between clinical measures and the observed profiles, the inclusion of more participant in the alternative experimental profiles (*Atypical representation – Low IN* and *High IN)* is required, considering, however, that in this study more participants were included than in previous studies of reverse correlations carried in the auditory modality (30 participants in each group versus 10 to 20 in other studies) (Adl Zarrabi et al., 2024; Goupil et al., 2021; Ponsot, Arias et al., 2018; Ponsot, Burred et al., 2018). However, due to the small sample size in the current study, the lack of a difference between groups may be attributed to insufficient statistical power. Future research should focus on increasing the sample size to improve the ability to detect effects with greater reliability.

Because the reverse correlation task was a part of a larger protocol, the addition of another task was not possible and thus, no measure of emotional prosody production has been performed. Thought it could have informed on the implication of different representation of vocal emotion on the production. In fact, it has been shown that the measurement of certain acoustic indices during prosodic productions (frequency of fundamental and formants and mean Harmonic to Noise Ratio, among others) enables a precise and sensitive classification according to a diagnosis of autism in children (Briend et al., 2023). Combining this information with the acoustic indices of the mental representation of vocal smile could enable this classificatory analysis to be extended. Furthermore, including a measure of emotional prosody processing in this study could provide valuable insights into the challenges faced by autistic individuals. While the prosodic difficulties in our ASD group are clear and clinically significant, they are more challenging to assess using standardized measurement tools.

Musical practice has been shown to enable a finer analysis of the acoustic signal (Molnar-Szakacs & Heaton, 2012), and this could have a beneficial effect on the development of vocal emotion representation in autism (Redondo Pedregal & Heaton, 2021). This finer acoustic analysis might play a role in the analysis of sounds with added social content. Unfortunately, such a detailed analysis of musical practice was not possible in the tested sample due to an insufficient number of musician participants (11/29 in the NT and 8/28 in the ASD group). In the future, it would be interesting to include more musician participants in order to test these potentially beneficial effects on perceptive representations of vocal emotions.

5. Conclusion

To summarize, this study allowed to estimate mental representation of vocal smile, and the internal noise level associated in a clinical population with a short and simple paradigm highlighting preserved processes in autistic adults. This study however revealed different profiles according to both mental representation and internal noise that could contribute to difficulties in recognition and response to emotional prosody in some autistic individuals. Measurement of emotional contagion evoked by prosodic voices in autism should allow to go further in the understanding of prosody atypicalities in the autistic participants who display typical and robust perceptive representation.

Author's contributions

Annabelle Merchie, Jean-Julien Aucouturier and Marie Gomot designed the study. Emmanuelle Houy-Durand was responsible for the recruitment and the clinical assessment of patients. Annabelle Merchie and Zoé Ranty performed data acquisition. Annabelle Merchie, Zoé Ranty, Aynaz Adl Zarrabi, Jean-Julien Aucouturier and Marie Gomot were responsible for data and statistical analyses. Annabelle Merchie, Jean-Julien Aucouturier and Marie Gomot wrote the first version of the manuscript. All authors were involved in preparing and reviewing the manuscript.

CRediT authorship contribution statement

Merchie Annabelle: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Aucouturier Jean-Julien:** Visualization, Validation, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Gomot Marie:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization. **Ranty Zoé :** Writing – review & editing, Investigation, Formal analysis, Data curation. **Adl Zarrabi Aynaz:**

Validation, Software. **Bonnet-Brilhault Frédérique:** Supervision. **Houy-Durand Emmanuelle:** Supervision, Investigation.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The protocol received approval from Ethics Committee (PROSCEA2017/23; ID RCB: 2017-A00756–47).

Funding

This work was supported by the INSERM (National Institute for Health and Medical Research). Annabelle Merchie was funded by the FEANS (Fondation Européenne pour l'Avancement des Neurosciences) and the Agence Nationale de la Recherche (ANR-19-CE37–0022-01). The funding sources were not involved in the study design, data collection and analysis, or the writing of the report.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Annabelle Merchie reports financial support was provided by Agence Nationale de la Recherche. Annabelle Merchie reports financial support was provided by Fondation Européenne pour l'avancement des neurosciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Adl Zarrabi, A., Jeulin, M., Bardet, P., Commère, P., Naccache, L., Aucouturier, J.-J., Ponsot, E., & Villain, M. (2024). A simple psychophysical procedure separates representational and noise components in impairments of speech prosody perception after right-hemisphere stroke. *Scientific Reports, 14*(1), Article 15194. https://doi.org/10.1038/s41598-024-64295-y

Andronikof, A., & Fontan, P. (2016). Grounia Efimovna Soukhareva: La première description du syndrome dit d'Asperger. *Neuropsychiatrie Dèleľött l'Enfance et Dèleľött l'Adolescence, 64*(1), 58–70. https://doi.org/10.1016/j.neurenf.2015.07.007

Arias, P., Belin, P., & Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. *Current Biology, 28*(14), R782–R783. https://doi.org/10.1016/j.cub.2018.05.084

Arias Sarah, P., Hall, L., Saitovitch, A., Aucouturier, J.-J., Zilbovicius, M., & Johansson, P. (2023). Pupil dilation reflects the dynamic integration of audiovisual emotional speech. Article 1. *Scientific Reports, 13*(1). https://doi.org/10.1038/s41598-023-32133-2

Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Communication, 40*(1–2), 87–97. https://doi.org/10.1016/S0167-6393(02)00077-8

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Beall, P. M., Moody, E. J., McIntosh, D. N., Hepburn, S. L., & Reed, C. L. (2008). Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *Journal of Experimental Child Psychology, 101*(3), 206–223. https://doi.org/10.1016/j.jecp.2008.04.004

Black, M. H., Chen, N. T. M., Iyer, K. K., Lipp, O. V., Bölte, S., Falkmer, M., Tan, T., & Girdler, S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neuroscience & Biobehavioral Reviews, 80*, 488–515. https://doi.org/10.1016/j.neubiorev.2017.06.016

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., Burack, J. A., & Mottron, L. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia, 48*(9), 2465–2475. https://doi.org/10.1016/j.neuropsychologia.2010.04.020

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *J Cogn Neurosci, 15*, 226–235.

Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders, 5*(4), 1567–1573. https://doi.org/10.1016/j.rasd.2011.03.002

Briend, F., David, C., Silleresi, S., Malvy, J., Ferré, S., & Latinus, M. (2023). Voice acoustics allow classifying autism spectrum disorder with high accuracy. *Translational Psychiatry, 13*(1), 1–8. https://doi.org/10.1038/s41398-023-02554-8

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology, 28*(1), 333–361. https://doi.org/10.1080/10463283.2017.1381469

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. *Trends in Cognitive Sciences, 16*(12), 573–574. https://doi.org/10.1016/j.tics.2012.10.005

Brooks, P. J., & Ploog, B. O. (2013). Attention to emotional tone of voice in speech perception in children with autism. *Research in Autism Spectrum Disorders, 7*(7), 845–857. https://doi.org/10.1016/j.rasd.2013.03.003

Burred, J. J., Ponsot, E., Goupil, L., Liuni, M., & Aucouturier, J.-J. (2019). CLEESE: An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLOS ONE, 14*(4), Article e0205943. https://doi.org/10.1371/journal.pone.0205943

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*, 1–36. https://doi.org/10.18637/jss.v061.i06

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia, 49*(3), 507–517. https://doi.org/10.1016/j.neuropsychologia.2010.11.042

Davis, G., & Plaisted-Grant, K. (2015). Low endogenous neural noise in autism. *Autism: The International Journal of Research and Practice, 19*(3), 351–362. https://doi.org/10.1177/1362361314552198

Day, T. C., Malik, I., Boateng, S., Hauschild, K. M., & Lerner, M. D. (2023). Vocal emotion recognition in Autism: Behavioral performance and event-related potential (ERP) response. *Journal of Autism and Developmental Disorders, 54*(4), 1235–1248. https://doi.org/10.1007/s10803-023-05898-8

Dawson, G., Hill, D., Spencer, A., Galpert, L., & Watson, L. (1990). Affective exchanges between young autistic children and their mothers. *Journal of Abnormal Child Psychology, 18*(3), 335–345. https://doi.org/10.1007/BF00916569

de Gelder, B., Vroomen, J., & Van Der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology, 3*(1), 69–86. https://doi.org/10.1080/09541449108406220

Ewing, L., Pellicano, E., King, H., Lennuyeux-Comnene, L., Farran, E. K., Karmiloff-Smith, A., & Smith, M. L. (2018). Atypical information-use in children with autism spectrum disorder during judgments of child and adult face identity. *Developmental Neuropsychology, 43*(4), 370–384. https://doi.org/10.1080/87565641.2018.1449846

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. Article 4. *Nature Reviews Neuroscience, 9*(4). https://doi.org/10.1038/nrn2258

Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications, 12*(1). https://doi.org/10.1038/s41467-020-20649-4

Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *J Speech Lang Hear Res JSLHR, 53*, 778–793.

Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder help improve clinical practice? *Frontiers in Psychiatry, 7*. ⟨https://www.frontiersin.org/articles/10.3389/fpsyt.2016.00107⟩.

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology, 25*(6), 771–782. https://doi.org/10.1080/02643290802336277

Hoffmann, J., Travers-Podmaniczky, G., Pelzl, M. A., Brück, C., Jacob, H., Hölz, L., Martinelli, A., & Wildgruber, D. (2023). Impairments in recognition of emotional facial expressions, affective prosody, and multisensory facilitation of response time in high-functioning autism. *Frontiers in Psychiatry, 14*. https://doi.org/10.3389/fpsyt.2023.1151665

Hubbard, K., & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research, 36*(2), 159–173. https://doi.org/10.1007/s10936-006-9037-4

Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder: Affective prosody in ASD. *Autism Research, 10*(12), 1991–2001. https://doi.org/10.1002/aur.1847

Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders, 36*(1), 77–90. https://doi.org/10.1007/s10803-005-0044-3

Jertberg, R. M., Begeer, S., Geurts, H. M., Chakrabarti, B., & Van der Burg, E. (2024). Age, not autism, influences multisensory integration of speech stimuli among adults in a McGurk/MacDonald paradigm. *European Journal of Neuroscience, 59*(11), 2979–2994. https://doi.org/10.1111/ejn.16319

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kassambara, A. (2020). ggpubr: "ggplot2" Based Publication Ready Plots. ⟨https://CRAN.R-project.org/package=ggpubr⟩.

Kassambara, A., & Mundt, F. (2017). Package 'factoextra.'. *Extract and Visualize the Results of Multivariate Data Analyses, 76*(2).

Kern, J. K., Garver, C. R., Carmody, T., Andrews, A. A., Trivedi, M. H., & Mehta, J. A. (2007). Examining sensory quadrants in autism. *Research in Autism Spectrum Disorders, 1*(2), 185–193. https://doi.org/10.1016/j.rasd.2006.09.002

Kissine, M., Bertels, J., Deconinck, N., Passeri, G., & Deliens, G. (2021). Audio-visual integration in nonverbal or minimally verbal young autistic children. *Journal of Experimental Psychology: General, 150*(10), 2137–2157. https://doi.org/10.1037/xge0001040

Knoblauch, K., & Maloney, L. T. (2008). Estimating classification images with generalized linear and additive models. *Journal of Vision, 8*(16), 10. https://doi.org/10.1167/8.16.10

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00302

Le Sourn-Bissaoui, S., Aguert, M., Girard, P., Chevreuil, C., & Laval, V. (2013). Emotional speech comprehension in children and adolescents with autism spectrum disorders. *Journal of Communication Disorders, 46*(4), 309–320. https://doi.org/10.1016/j.jcomdis.2013.03.002

Leung, F. Y. N., Sin, J., Dawson, C., Ong, J. H., Zhao, C., Veić, A., & Liu, F. (2022). Emotion recognition across visual and auditory modalities in autism spectrum disorder: A systematic review and meta-analysis. *Developmental Review, 63*, Article 101000. https://doi.org/10.1016/j.dr.2021.101000

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685. https://doi.org/10.1007/BF02172145

Magnuson, J. R., Iarocci, G., Doesburg, S. M., & Moreno, S. (2020). Increased intra-subject variability of reaction times and single-trial event-related potential components in children with autism spectrum disorder. *Autism Research, 13*(2), 221–229. https://doi.org/10.1002/aur.2210

Manenti, M., Ferré, S., Tuller, L., Houy-Durand, E., Bonnet-Brilhault, F., & Prévost, P. (2024). Profiles of structural language and nonverbal intellectual abilities in verbal autistic adults. *Research in Autism Spectrum Disorders, 114*, Article 102361. https://doi.org/10.1016/j.rasd.2024.102361

Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science, 28*(2), 209–226. https://doi.org/10.1207/s15516709cog2802_4

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders, 38*(4), 325–350. https://doi.org/10.1080/1368282031000154204

Merchie, A., Ranty, Z., Aguillon-Hernandez, N., Aucouturier, J.-J., Wardak, C., & Gomot, M. (2024). Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance. *Scientific Reports, 14*(1), Article 25043. https://doi.org/10.1038/s41598-024-74848-w

Miller, L. K. (1999). The Savant Syndrome: Intellectual impairment and exceptional skill. *Psychological Bulletin, 125*(1), 31–46. https://doi.org/10.1037/0033-2909.125.1.31

Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences, 1252*(1), 318–324. https://doi.org/10.1111/j.1749-6632.2012.06465.x

Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review, 17*(6), 802–808. https://doi.org/10.3758/PBR.17.6.802

Neri, P. (2011). Coarse to fine dynamics of monocular and binocular processing in human pattern vision. *Proceedings of the National Academy of Sciences, 108*(26), 10726–10731. https://doi.org/10.1073/pnas.1101246108

Park, W. J., Schauder, K. B., Zhang, R., Bennetto, L., & Tadin, D. (2017). High internal noise and poor external noise filtering characterize perception in autism spectrum disorder. Article 1. *Scientific Reports, 7*(1). https://doi.org/10.1038/s41598-017-17676-5

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220. https://doi.org/10.1007/s10803-004-1999-1

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research, 50*(4), 1015–1028. https://doi.org/10.1044/1092-4388(2007/071)

Ponsot, E., Arias, P., & Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. *The Journal of the Acoustical Society of America, 143*(1), EL19–EL24. https://doi.org/10.1121/1.5020989

Ponsot, E., Burred, J. J., Belin, P., & Aucouturier, J.-J. (2018). Cracking the social code of speech prosody using reverse correlation. *Proceedings of the National Academy of Sciences, 115*(15), 3972–3977. https://doi.org/10.1073/pnas.1716090115

Redondo Pedregal, C., & Heaton, P. (2021). Autism, music and Alexithymia: A musical intervention to enhance emotion recognition in adolescents with ASD. *Research in Developmental Disabilities, 116*, Article 104040. https://doi.org/10.1016/j.ridd.2021.104040

Righi, G., Tenenbaum, E. J., McCormick, C., Blossom, M., Amso, D., & Sheinkopf, S. J. (2018). Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. *Autism Research, 11*(4), 645–653. https://doi.org/10.1002/aur.1918

Rimland, B., & Fein, D. (1988). Special talents of autistic savants. *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 474–492). The Guilford Press.

Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders, 43*(4), 775–784. https://doi.org/10.1007/s10803-012-1608-7

Scheerer, N. E., Shafai, F., Stevenson, R. A., & Iarocci, G. (2020). Affective prosody perception and the relation to social competence in autistic and typically developing children. *Journal of Abnormal Child Psychology, 48*(7), 965–975. https://doi.org/10.1007/s10802-020-00644-5

Song, Y., Kawabe, T., Hakoda, Y., & Du, X. (2012). Do the eyes have it? Extraction of identity and positive expression from another's eyes in autism, probed using "Bubbles. *Brain and Development, 34*(7), 584–590. https://doi.org/10.1016/j.braindev.2011.09.009

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics, 27*(1), 24–27. https://doi.org/10.3758/BF03199901

Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *The Journal of the Acoustical Society of America, 96*(4), 2101–2107. https://doi.org/10.1121/1.410151

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language & Communication Disorders, 50*(4), 452–466. https://doi.org/10.1111/1460-6984.12146

Team, R.C. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. ⟨https://www.R-project.org/⟩.

Team, Rs. (2020). RStudio: Integrated Development for R. In RStudio, PBC. ⟨http://www.rstudio.com/⟩.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review, 121*(4), 649. https://doi.org/10.1037/a0037665

Van de Cruys, S., Van der Hallen, R., & Wagemans, J. (2017). Disentangling signal and noise in autism spectrum disorder. *Brain and Cognition, 112*, 78–83. https://doi.org/10.1016/j.bandc.2016.08.004

Vilidaite, G., Yu, M., & Baker, D. H. (2017). Internal noise estimates correlate with autistic traits. *Autism Research: Official Journal of the International Society for Autism Research, 10*(8), 1384–1391. https://doi.org/10.1002/aur.1781

Wang, J.-E., & Tsao, F.-M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities, 37*, 162–170. https://doi.org/10.1016/j.ridd.2014.11.013

Wang, L., Ong, J. H., Ponsot, E., Hou, Q., Jiang, C., & Liu, F. (2022). Mental representations of speech and musical pitch contours reveal a diversity of profiles in autism spectrum disorder. *Autism*. , Article 13623613221111207. https://doi.org/10.1177/13623613221111207

Ward, J. (2019). Individual differences in sensory sensitivity: A synthesizing framework and evidence from normal variation and developmental conditions. *Cognitive Neuroscience, 10*(3), 139–157. https://doi.org/10.1080/17588928.2018.1557131

Wechsler, D. (2012). Wechsler Adult Intelligence Scale—Fourth Edition [Dataset]. https://doi.org/10.1037/t15169–000.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. ⟨https://ggplot2.tidyverse.org⟩.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2021). dplyr: A Grammar of Data Manipulation. ⟨https://CRAN.R-project.org/package=dplyr⟩.

Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the face: Sensorimotor simulation contributes to facial expression recognition. *Trends in Cognitive Sciences, 20*(3), 227–240. https://doi.org/10.1016/j.tics.2015.12.010

Zhang, M., Xu, S., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2022). Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism, 26*(4), 798–813. https://doi.org/10.1177/1362361321995725