

Improving emotional speech processing in autism spectrum disorders: toward the elaboration of a drug-free intervention based on Social Stories™ and NAO social robot interactions

Submission date 11/09/2020	Recruitment status No longer recruiting	<input checked="" type="checkbox"/> Prospectively registered <input checked="" type="checkbox"/> Protocol
Registration date 17/09/2020	Overall study status Completed	<input type="checkbox"/> Statistical analysis plan <input checked="" type="checkbox"/> Results
Last Edited 19/03/2024	Condition category Mental and Behavioural Disorders	<input checked="" type="checkbox"/> Individual participant data

Plain English summary of protocol

Background and study aims

Autism Spectrum Disorders (ASD) are neurodevelopmental disorders that are characterized by social and emotional deficits. Therefore, autistic people may have general difficulties behaving in a social context, expressing their emotions appropriately and inferring others' minds. Particularly, autistic people may be impaired at distinguishing between emotional intonations in voices during conversations which can cause misunderstandings and negatively affect their social life. Drug-free training based on interactions with a humanoid social robot has been recently developed and was found to help autistic children improve their social abilities. Consequently, the aim of this study is to develop a new kind of social robot-based training using NAO, which is a talking and friendly-shaped 57-cm-high humanoid robot, to help autistic children distinguish emotion in voices. As autistic people are particularly interested in technological devices and digital environments, synthetic voices will be used to improve the learning and gradually reach the understanding of human voices. As the participants are Mexican and emotional expression differs between cultures, the researchers need to determine acoustic cues in human voices that allow the differentiation of emotional intonations as uttered in a Mexican way. The results are compared to those obtained from a Castilian Spanish database. Once the researchers know how to differentiate emotions in voices, they will be able to correctly digitally synthesize emotional utterances that approximate human voices at a greater or lesser extent. Finally, those newly acquired synthesized voices will be used in robot-based training for improving emotional discrimination abilities in autistic children.

Who can participate?

Experiment 1: typically developed adults (aged 19 to 35) and children (aged 9 to 11)

Experiments 2, 3, and 4: autistic and typically developed children aged 9 to 11

What does the study involve?

Participants from the first experiment will be typically developed adults and children who will have to utter words with a particular emotional intonation (e.g. anger, sadness) into a microphone. Sessions will last 1 hour for adults and 40 minutes for children. Participants from the second experiment will be typically developed and autistic children who will have to listen to the emotional utterances recorded in experiment 1 while answering an attentional task on a computer and their brain's activity is recorded by electrodes on their scalp (electroencephalography). Participants from the third experiment will listen to both human and synthesized voices newly created from the recordings while their brain's activity is recorded by electroencephalography. Finally, the synthesized voice screened in experiment 3 and the human voices will be used in a training based on interactions with the NAO robot which will be programmed to tell interactive stories (Social Stories™) including relevant cues to improve the understanding of emotional intonation. Participants will be autistic children. Typically developed children will participate only in assessment sessions (baseline, mid-training, end of training and two follow-up sessions where emotion understanding will be evaluated). The training will be 15-session long where sessions 1, 7, 13, 14 and 15 will be assessment sessions and sessions 2 to 13 will be for training. Therefore, sessions 7 and 13 will include both assessment and training. Sessions 14 and 15 will be 2 and 4 weeks after the end of the training, respectively. Each assessment will last 35 minutes and training lasts 10 minutes. Parents will have to answer the Autism Spectrum Rating Scale (20 minutes) at sessions 1, 14 and 15.

What are the possible benefits and risks of participating?

There will be no risk in any of the four experiments. Participants from experiments 1, 2, and 3 will respectively help to promote the understanding of cross-cultural differences in emotional intonations and of brain activity correlates of emotion processing in autistic and typically developed children. Experiment 4 aims to improve the distinction of emotional intonations in voice for autistic children.

Where is the study run from?

Tecnologico de Monterrey, Monterrey (Mexico)

When is the study starting and how long is it expected to run for?

February 2020 to June 2023

Who is funding the study?

National Council of Science and Technology (Mexico)

Who is the main contact?

Mathilde Marie Duville

A00829725@itesm.mx

Contact information

Type(s)

Scientific

Contact name

Miss Mathilde Marie Duville

ORCID ID

<https://orcid.org/0000-0001-7309-3541>

Contact details

Tecnologico de Monterrey
Ave. Eugenio Garza Sada 2501
Monterrey
Mexico
64849
+81 (0)34 11 34 13
A00829725@itesm.mx

Additional identifiers**Clinical Trials Information System (CTIS)**

Nil known

ClinicalTrials.gov (NCT)

Nil known

Protocol serial number

CONBIOETICA19CEI01120161017

Study information**Scientific Title**

Electroencephalographic correlate of Mexican Spanish emotional speech processing in Autism Spectrum Disorders: to a Social Story™ and robot-based intervention

Study objectives

Current study hypothesis as of 10/10/2022:

1. Cross-cultural differences between Castilian and Mexican Spanish in emotional prosody can be highlighted by analyzing duration, intensity, and spectral characteristics of emotional speech signals.
2. Autistic electroencephalographic correlates of emotional speech processing during concurrent visual task is impaired and less affected by attentional load compared to typically-developed children.
3. Autistic electroencephalographic correlates of emotional speech processing approximate non-pathological ones when speech is conveyed by synthesized voices.
4. A drug-free social story and robot-based intervention that implements the use of synthesized voices results in improvement of emotional prosodies differentiation skills in autistic children.

Previous study hypothesis:

1. Cross-cultural differences between Castilian and Mexican Spanish in emotional prosody can be highlighted by analyzing duration, intensity and spectral characteristics of emotional speech signals.
2. Autistic electroencephalographic and behavioural correlates of emotional speech processing approximate non-pathological ones when speech is conveyed by synthesized voices.
3. A drug-free social story and robot-based intervention that implements the use of synthesized voices results in improvement of emotional prosodies differentiation skills in autistic children.

Ethics approval required

Old ethics approval format

Ethics approval(s)

Approved 14/07/2020, Ethical Committee of the School of Medicine of Tecnológico de Monterrey (Av. Ignacio Morones Prieto 3000 Pte., Col. Los Doctores, CP: 64710, Monterrey, Nuevo León, México; +81 (0)88 88 21 07; ana.villalpando@itesm.mx), ref: autismoEEG2020

Study design

Experiment 1: Observational single-center analytical case-control retrospective study

Experiment 2: Observational single-center analytical case-control retrospective study

Added 10/10/2022: Experiment 3: Observational single-center analytical case-control retrospective study

Experiment 4: Interventional single-center analytical prospective randomized controlled trial

Primary study design

Interventional

Study type(s)

Quality of life

Health condition(s) or problem(s) studied

Autism spectrum disorders

Interventions

Current interventions as of 10/10/2022:

Experiment 1: observational trials – non-pathological participants

Speech data will be recorded from 8 non-professional volunteer actors: 2 adult males and females, 2 children male and female recruited from a public announcement posted on Instituto Tecnológico de Monterrey, Campus Monterrey. Participants will be asked to read and get a grip on the entire words' dataset and will be given as much time as he/she needs. Then he/she will be instructed to read each word one time with an emotional intonation: neutral, happy, angry, sad, disgusted, or afraid. All words from a particular emotion will be uttered successively to facilitate this process. Participants will be asked to wait at least 5 seconds between two utterances to ensure a basal emotional state before each recording. Written informed consents will be obtained from all participants at the beginning of the session. Each session will last approximately 1 hour for adults and 40 minutes for children.

Experiment 2: observational trials – high-functioning autistic and non-pathological children aged 9-11

The task will proceed in a regular office room where the participant will be comfortably seated in a chair in front of a computer screen. Children will be asked to 1) view neutral images, 2) answer a moderate attentional load Multiple Object Tracking (MOT) task, 3) answer a high attentional load MOT task while emotional stimuli uttered with the human voices recorded in Experiment 1 will be displayed. The stimuli will be presented at 60 dB via the Shure SRH1840 audio headset that has a flat response frequency response to accurately reproduce the input audio signal. The stimulus onset asynchrony will be of 3,110 ms. Task and stimuli sequences will be randomly distributed and will differ for every child. Instructions will be written on the screen and explained verbally to ensure correct understanding before starting the task. Written informed consents will be obtained from all participants and their legal tutors before starting the session.

Experiment 3: observational trials – high-functioning autistic and non-pathological children aged 9-11

The task will proceed in a regular office room where the participant will be comfortably seated in a chair in front of a computer screen while answering a MOT task and will be instructed to pay no attention to the auditive stimuli. Before starting the experiment, the EEG cap will be mounted on the participant's head and the equipment for electroencephalographic activity recording will be set up. Once the participant is ready and calm, the experimental task will start. Emotional and neutral words uttered by human voices recorded in Experiment 1 and synthesized voices created in Experiment 2 will be randomly displayed to the participant. To facilitate the comparison of naturalness levels, six stimulus blocks will be created, each containing 600 stimuli corresponding to one of the five levels, or the human voice. Each block will consist of 480 frequently presented neutral-uttered standard stimuli (80%), and occasional 24 anger (4%), 24 disgust (4%), 24 fear (4%), 24 happiness (4%), and 24 sadness (4%) deviant stimuli. Standard and deviant stimuli will be equally uttered with a male, woman, or child voice. The stimuli will be presented pseudo-randomly with deviant stimuli being preceded by at least two standard stimuli. Each participant will listen to three blocks for an approximatively 30-min task. The stimulus sequences for each block will be fixed, and the order of blocks will be counterbalanced and randomized between participants. The stimulus onset asynchrony will be of 1,300 ms. The stimuli will be presented at 60 dB via the Shure SRH1840 audio headset that has a flat response frequency response to accurately reproduce the input audio signal. Autistic participants will be accompanied by their parent if necessary. Instructions will be written on the screen and explained verbally to ensure correct understanding before starting the task. Written informed consents will be obtained from all participants and their legal tutors before starting the session.

Experiment 4: interventional trials – high-functioning autistic and non-pathological children aged 9-11

Typically developed participants will be recruited from a public announcement posted on Tecnologico de Monterrey, Campus Monterrey. Autistic participants will be recruited by contacting parents of autistic children attending San Jose Hospital from Tecnologico de Monterrey.

The intervention will consist of six-time per week sessions for 2 weeks conducted in a one-to-one format. Each session will last approximately ten minutes and will be composed of two Social Stories™. Presentation order will be random and different for each participant. The stories will be uttered by the NAO social robot that will be programmed using Choregraphe software. Emotional and neutral speech will be uttered using the synthesized voiced selected in Experiment 2. Each Social Story™ will be written to train emotion differentiation and recognition through prosody. Questions about the emotion conveyed will be asked. Correct responses will lead to a social reinforcement (e.g.: "Well done!" *in Spanish* with a happy prosody and synthesized voice). Incorrect or absence of response will result in repeating the part of the story that contains the information necessary to answer the question. Then the question will be asked again, and a correct answer will lead to a social reinforcement. In case of an incorrect answer, the correct answer will be told by the robot. In case of correct answers, the robot will specify that in daily life, this emotion may not be uttered by the synthesized voice but will be better uttered by his entourage with a more natural voice. Then the robot will repeat the same emotional utterance with the human voice and ask which emotion is conveyed. Again, correct responses will lead to social reinforcement. Incorrect or absence of response will result in repeating the part of the story that contains the information necessary to answer the question. Then the question will be asked again, and a correct answer will lead to a social reinforcement. In case of an incorrect answer, the correct answer will be told by the robot. Each session will start with robot NAO inviting the participant to play the "emotion game" and explaining the instructions. The robot's movements and eye lighting will only be programmed for this part. During Social Stories™, NAO will remain still to focus the child's attention to the audio

information. All session long, NAO will be remotely controlled by the researcher with a Wizard-Of-Oz method using the Graphical User Interface (GUI) provided by the system. Written informed consent will be obtained from all participants and their legal tutors before starting the intervention. A total of 20 participants will take part in this experiment and will be divided into four groups:

Group 1: autistic children receiving robot, synthesized and human voices intervention (5 participants)

Group 2: autistic children receiving robot and human voice intervention (5 participants)

Group 3: autistic children not receiving the intervention (5 participants)

Group 4: typically developed children not receiving the intervention (5 participants)

Group 1 will be the experimental group and groups 2 to 4 will be controls. Children from group 2 will receive the same intervention as children from group 1, except for the use of human voices only. This group will help to evaluate the efficacy of the use of synthesized voices in the intervention. Autistic participants will be randomly assigned to any of groups 1, 2 or 3 (using Excel random function). All four groups will be included in pre- (1st session), mid (7th session), end (13th session) and follow-up assessments (2 and 6 weeks after the end of the intervention). Please refer to outcome measures for more details about the assessments.

Previous interventions:

Experiment 1: observational trials – non-pathological participants

Speech data will be recorded from 8 non-professional volunteer actors: 2 adult males and females, 2 children male and female recruited from a public announcement posted on Instituto Tecnológico de Monterrey, Campus Monterrey. Participants will be asked to read and get a grip on the entire words' dataset and will be given as much time as he/she needs. Then he/she will be instructed to read each word one time with an emotional intonation: neutral, happy, angry, sad, disgusted or afraid. All words from a particular emotion will be uttered successively in order to facilitate this process. Participants will be asked to wait at least 5 seconds between two utterances in order to ensure a basal emotional state before each recording. Written informed consents will be obtained from all participants at the beginning of the session. Each session will last approximatively 1 hour for adults and 40 minutes for children.

Experiment 2: observational trials – high-functioning autistic and non-pathological children aged 9-11

The task will proceed in a regular office room where the participant will be comfortably seated in a chair in front of a computer screen. Before starting the experiment, the EEG cap will be mounted on the participant's head and the equipment for electroencephalographic activity recording will be set up. Once the participant is ready and calm, the experimental task will start. Emotional and neutral words uttered by human voices recorded in Experiment 1 and synthesized voices created in Experiment 2 will be randomly displayed to the participant. After each utterance, neutral, happy, angry, sad, disgusted and feared face emojis will appear on the computer screen, and the participant will have to choose which one fits best with the emotion conveyed by the word according to its prosody. Utterances will be displayed in an audio headset at 60 dBA. They will be spaced with an 8-second interval in order to come back to a basal emotional state before each word. For an approximative 35-min task, about 24 stimuli per emotion (anger, disgust, fear, neutral, sadness and happiness) will be presented, for a total of 144 stimuli. Before starting the actual paradigm and after the explanation of the task, trial stimuli will be displayed (one for each emotion). The participant will be informed that he/she will not have to answer these trials. The utterance will be displayed in the audio headset and the correct response will automatically be displayed on the screen by encircling the correct emoji. Right after, the participant will be informed that for the following stimuli he/she will have to answer. Then the task will start. Instructions will be written on the screen and explained verbally

to ensure correct understanding before starting the task. Written informed consents will be obtained from all participants and their legal tutors before starting the session.

Experiment 3: interventional trials – high-functioning autistic and non-pathological children aged 9-11

Typically-developed participants will be recruited from a public announcement posted on Tecnológico de Monterrey, Campus Monterrey. Autistic participants will be recruited by contacting parents of autistic children attending San Jose Hospital from Tecnológico de Monterrey.

The intervention will consist of six-time per week sessions for 2 weeks conducted in a one-to-one format. Each session will last approximately ten minutes and will be composed of two Social Stories™. Presentation order will be random and different for each participant. The stories will be uttered by the NAO social robot that will be programmed using Choregraphe software. Emotional and neutral speech will be uttered using the synthesized voiced selected in Experiment 2. Each Social Story™ will be written to train emotion differentiation and recognition through prosody. Questions about the emotion conveyed will be asked. Correct responses will lead to a social reinforcement (e.g.: “Well done!” *in Spanish* with a happy prosody and synthesized voice). Incorrect or absence of response will result in repeating the part of the story that contains the information necessary to answer the question. Then the question will be asked again, and a correct answer will lead to a social reinforcement. In case of an incorrect answer, the correct answer will be told by the robot. In case of correct answers, the robot will specify that in daily life, this emotion may not be uttered by the synthesized voice but will be better uttered by his entourage with a more natural voice. Then the robot will repeat the same emotional utterance with the human voice and ask which emotion is conveyed. Again, correct responses will lead to social reinforcement. Incorrect or absence of response will result in repeating the part of the story that contains the information necessary to answer the question. Then the question will be asked again, and a correct answer will lead to a social reinforcement. In case of an incorrect answer, the correct answer will be told by the robot. Each session will start with robot NAO inviting the participant to play the “emotion game” and explaining the instructions. The robot’s movements and eye lighting will only be programmed for this part. During Social Stories™, NAO will remain still in order to focus the child’s attention to the audio information. All session long, NAO will be remotely controlled by the researcher with a Wizard-Of-Oz method using the Graphical User Interface (GUI) provided by the system. Written informed consent will be obtained from all participants and their legal tutors before starting the intervention. A total of 20 participants will take part in this experiment and will be divided into four groups:

Group 1: autistic children receiving robot, synthesized and human voices intervention (5 participants)

Group 2: autistic children receiving robot and human voice intervention (5 participants)

Group 3: autistic children not receiving the intervention (5 participants)

Group 4: typically-developed children not receiving the intervention (5 participants)

Group 1 will be the experimental group and groups 2 to 4 will be controls. Children from group 2 will receive the same intervention as children from group 1, with the exception of the use of human voices only. This group will help to evaluate the efficacy of the use of synthesized voices in the intervention. Autistic participants will be randomly assigned to any of groups 1, 2 or 3 (using Excel random function). All four groups will be included in pre- (1st session), mid (7th session), end (13th session) and follow-up assessments (2 and 6 weeks after the end of the intervention). Please refer to outcome measures for more details about the assessments.

Intervention Type

Behavioural

Primary outcome(s)

Current primary outcome measures as of 10/10/2022:

1. Autistic symptomatology measured by the Autism Spectrum Rating Scales. Parents of both groups will answer this questionnaire at pre-intervention (1st session) and follow-up sessions (2 and 6 weeks after the end of the intervention).
2. Electroencephalographic correlates and behavioural responses to a task related to emotional prosody processing measured with a 24-channel EEG system (Greentek Gelfree S3, mBrain Train), digitized and amplified at a sampling rate of 500Hz (pass band: 0.1-100Hz). Electrodes placement will be in accordance with the international 10/20 system. Measures will be done at pre-intervention (1st session), mid and end sessions (sessions 7 and 13) and follow-up sessions (2 and 6 weeks after the end of the intervention).

The emotional prosody task will be as follows:

The task will proceed in a regular office room where the participant will be comfortably seated in a chair in front of a computer screen. Before starting the experiment, the EEG cap will be mounted on the participant's head and the equipment for electroencephalographic activity recording will be set up. Once the participant is ready and calm, the experimental task will start. Emotional and neutral words uttered by human voices recorded in Experiment 1 and synthesized voices created in Experiment 3 will be randomly displayed to the participant. After each utterance, neutral, happy, angry, sad, disgusted and feared face emojis will appear on the computer screen, and the participant will have to choose which one fits best with the emotion conveyed by the word according to its prosody. Utterances will be displayed in an audio headset at 60 dBA. They will be spaced with an 8-second interval to come back to a basal emotional state before each word. For an approximative 35-min task, about 24 stimuli per emotion (anger, disgust, fear, neutral, sadness and happiness) will be presented, for a total of 144 stimuli. Before starting the actual paradigm and after the explanation of the task, trial stimuli will be displayed (one for each emotion). The participant will be informed that he/she will not have to answer these trials. The utterance will be displayed in the audio headset and the correct response will automatically be displayed on the screen by encircling the correct emoji. Right after, the participant will be informed that for the following stimuli he/she will have to answer. Then the task will start. Instructions will be written on the screen and explained verbally to ensure correct understanding before starting the task. Written informed consent will be obtained from all participants and their legal tutors before starting the session.

Previous primary outcome measures:

1. Autistic symptomatology measured by the Autism Spectrum Rating Scales. Parents of both groups will answer this questionnaire at pre-intervention (1st session) and follow-up sessions (2 and 6 weeks after the end of the intervention).
2. Electroencephalographic correlates and behavioural responses to a task related to emotional prosody processing measured with a 24-channel EEG system (Greentek Gelfree S3, mBrain Train), digitized and amplified at a sampling rate of 500Hz (pass band: 0.1-100Hz). Electrodes placement will be in accordance with the international 10/20 system. Measures will be done at pre-intervention (1st session), mid and end sessions (sessions 7 and 13) and follow-up sessions (2 and 6 weeks after the end of the intervention). The emotional prosody task will be similar to the one described for Experiment 2.

Key secondary outcome(s)

Current secondary outcome measures as of 10/10/2022:

Experiment 1: cross-cultural differences between Mexican and Castilian Spanish of emotional

prosodies measured by statistical analysis of the following physical acoustic features:

- 1.1. Rate (number of syllables per second)
- 1.2. Mean, maximum, minimum, and standard deviation pitch in hertz
- 1.3. Jitter local and jitter ppq5 in percentages
- 1.4. Shimmer local and shimmer apq5 in percentages
- 1.5. Harmonic-to-noise-ratio in decibels
- 1.6. Root mean square energy in volts
- 1.7. Mean, maximum, minimum intensity in decibels
- 1.8. Mean first, second and third formants in hertz
- 1.9. Mean first, second and third formants bandwidths in hertz
- 1.10. 13 first Mel Frequency Cepstral Coefficients (MFCC)

Each participant will be recorded in one session.

Experiment 2: electroencephalographic correlates of emotional processing under attentional load modulations. Event-Related Spectral Perturbations will be extracted to assess the effect of the visual task on emotional prosody processing, and to compare responses to the different emotions.

Experiment 3: electroencephalographic correlates of discrimination of emotional prosodies uttered by synthesized and human voices measured by the analysis of the amplitude of the Late Positive Potential, P200, N300 components and the number of correct/incorrect responses. Each participant will be recorded in one session.

Previous secondary outcome measures:

Experiment 1: cross-cultural differences between Mexican and Castilian Spanish of emotional prosodies measured by statistical analysis of the following physical acoustic features:

- 1.1. Rate (number of syllables per second)
- 1.2. Mean, maximum, minimum, and standard deviation pitch in hertz
- 1.3. Jitter local and jitter ppq5 in percentages
- 1.4. Shimmer local and shimmer apq5 in percentages
- 1.5. Harmonic-to-noise-ratio in decibels
- 1.6. Root mean square energy in volts
- 1.7. Mean, maximum, minimum intensity in decibels
- 1.8. Mean first, second and third formants in hertz
- 1.9. Mean first, second and third formants bandwidths in hertz
- 1.10. 13 first Mel Frequency Cepstral Coefficients (MFCC)

Each participant will be recorded at one session.

Experiment 2: electroencephalographic and behavioural correlates of discrimination of emotional prosodies uttered by synthesized and human voices measured by the analysis of the amplitude of the Late Positive Potential, P200, N300 components and the number of correct /incorrect responses. Each participant will be recorded at one session.

Completion date

30/06/2023

Eligibility

Key inclusion criteria

Current inclusion criteria as of 10/10/2022:

Experiment 1:

1. Healthy male and female adults and children (age range for children: 9-11 years old, age range for adults: 19-35 years old).

2. Participants will have to have grown up in Mexico to ensure a Mexican way of conveying emotions

Experiment 2 and 3:

1. High-functioning autistic volunteers aged 9-11 years old
2. All participants will be diagnosed by a clinician according to DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) or ICD-10 (International Classification of Diseases and Related Health Problems – Tenth Edition) criteria
3. Autistic symptomatology will be assessed by the Autism Spectrum Rating Scales – parent report). A T-score greater than 60 (slightly elevated to very elevated scores for autistic behaviors) on Social/Communication, DSM-IV-TR (Text Revision), Peer and Adult Socialization and Social/Emotional Reciprocity domains will be considered as an inclusion criterion. Typically developed (TD) volunteers aged 9-11-year-old will also be included in this study. In the same way that for autistic children, Autism Spectrum Rating Scales (ASRS) will be used, and children with a score inferior to 60 on all scales will be included.

Experiment 4:

Autistic and TD participants involved in this study will have to meet the same criteria as those described in Experiment 2. However, participants from Experiment 2 will not be included in Experiment 3 to avoid any bias elicited by a former exposition to the synthetic voices under study.

Previous inclusion criteria:

Experiment 1:

1. Healthy male and female adults and children (age range for children: 9-11 years old, age range for adults: 19-35 years old).
2. Participants will have to have grown up in Mexico in order to ensure a Mexican way of conveying emotions

Experiment 2:

1. High-functioning autistic volunteers aged 9-11 years old
2. All participants will be diagnosed by a clinician according to DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) or ICD-10 (International Classification of Diseases and Related Health Problems – Tenth Edition) criteria
3. Autistic symptomatology will be assessed by the Autism Spectrum Rating Scales – parent report). A T-score greater than 60 (slightly elevated to very elevated scores for autistic behaviors) on Social/Communication, DSM-IV-TR (Text Revision), Peer and Adult Socialization and Social/Emotional Reciprocity domains will be considered as an inclusion criterion.

Typically-developed (TD) volunteers aged 9-11-year-old will also be included in this study. In the same way that for autistic children, Autism Spectrum Rating Scales (ASRS) will be used, and children with a score inferior to 60 in all scales will be included.

Experiment 3:

Autistic and TD participants involved in this study will have to meet the same criteria than those described in Experiment 2. However, participants from Experiment 2 will not be included in Experiment 3 in order to avoid any bias elicited by a former exposition to the synthetic voices under study.

Participant type(s)

Mixed

Healthy volunteers allowed

No

Age group

Mixed

Sex

All

Total final enrolment

80

Key exclusion criteria

Current exclusion criteria as of 10/10/2022:

Experiment 1:

1. Participants will be excluded if they have been diagnosed with any pathology that affects emotional behavior, hearing, or speech at the time of the study
2. Participants presenting sickness traits affecting voice timbre at the time of the study

Experiments 2 and 3:

Any TD or autistic participant will be excluded if:

1. She/he is under medication affecting the central or peripheral nervous system
2. She/he has hearing loss or deficits
3. She/he has a history of a developmental pathology and/or any disease affecting behavior and nervous system, apart from ASD for autistic participants

Experiment 4:

Autistic and TD participants involved in this study will have to meet the same criteria as those described in Experiment 2

Previous exclusion criteria:

Experiment 1:

1. Participants will be excluded if they have been diagnosed with any pathology that affects emotional behavior, hearing or speech at the time of the study
2. Participants presenting sickness traits affecting voice timbre at the time of the study

Experiment 2:

Any TD or autistic participant will be excluded if:

1. She/he is under medication affecting the central or peripheral nervous system
2. She/he has hearing loss or deficits
3. She/he has a history of a developmental pathology and/or any disease affecting behavior and nervous system, apart from ASD for autistic participants

Experiment 3:

Autistic and TD participants involved in this study will have to meet the same criteria than those described in Experiment 2

Date of first enrolment

01/10/2020

Date of final enrolment

28/02/2023

Locations

Countries of recruitment

Mexico

Study participating centre

Tecnologico de Monterrey

Ave. Eugenio Garza Sada 2501

Monterrey

Mexico

64849

Sponsor information

Organisation

Monterrey Institute of Technology and Higher Education

ROR

<https://ror.org/03ayjn504>

Funder(s)

Funder type

Government

Funder Name

Consejo Nacional de Ciencia y Tecnología (ref: 1061809)

Alternative Name(s)

Consejo Nacional de Humanidades, Ciencias y Tecnologías, Consejo Nacional de Ciencia y Tecnología, National Council of Humanities, Sciences and Technologies, Mexican National Council of Science and Technology, National Council for Science and Technology (CONACyT), National Council of Science and Technology, Mexico, Conahcyt

Funding Body Type

Government organisation

Funding Body Subtype

National government

Location
Mexico

Results and Publications

Individual participant data (IPD) sharing plan

The datasets generated during and/or analysed during the current study will be stored in a publically available repository. Human and synthesized voices generated from this project are openly available in Mendeley Data at <http://doi.org/10.17632/cy34mh68j9.5> (last access: 12/10/2022). EEG data generated from Experiment 2 are openly available in Mendeley Data at: <http://doi.org/10.17632/spwnt8t25y.1> (typically-developed children, last access on 12/10/2022), and at <http://doi.org/10.17632/7zf99hxxc9.1> (autistic children, last access on 12/10/2022). Data from studies 3 and 4 will be made available upon study completion.

IPD sharing plan summary

Stored in publicly available repository

Study outputs

Output type	Details	Date created	Date added	Peer reviewed?	Patient-facing?
Results article	protocol	06/12/2021	12/10/2022	Yes	No
Results article		14/03/2024	19/03/2024	Yes	No
Protocol article		26/02/2021	16/03/2021	Yes	No
Dataset	Participant information sheet	17/02/2022	12/10/2022	No	No
Dataset		16/08/2022	12/10/2022	No	No
Dataset		08/07/2022	12/10/2022	No	No
Participant information sheet			08/10/2020	No	Yes
Participant information sheet		11/11/2025	11/11/2025	No	Yes