cortex 155 (2022) 150–161

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

ELSEVIER

Behavioural Neurology

Emotion expression through spoken language in Huntington disease

Charlotte Gallezot [a,b,1], *Rachid Riad* [a,b,1], *Hadrien Titeux* [a], *Laurie Lemoine* [b,c], *Justine Montillot* [b,c], *Agnes Sliwinski* [b,c], *Jennifer Hamet Bagnou* [b,c], *Xuan Nga Cao* [a], *Katia Youssov* [b,c], *Emmanuel Dupoux* [a] and *Anne-Catherine Bachoud Levi* [b,c,*]

[a] *CoML/ENS/CNRS/EHESS/INRIA/PSL Research University, Paris, France*
[b] *NPI/ENS/INSERM U855/UPEC/PSL Research University, Creteil, France*
[c] *Assistance Publique-Hôpitaux de Paris, National Reference Center for Huntington's Disease, Groupe Henri-Mondor Albert-Chenevier, Créteil, France*

ARTICLE INFO

Article history:
Received 11 March 2022
Reviewed 11 April 2022
Revised 18 May 2022
Accepted 29 May 2022
Action editor: Stefano Cappa
Published online 19 July 2022

Keywords:
Huntington's disease
Speech
Machine learning
Emotions

ABSTRACT

Patients with Huntington's disease suffer from disturbances in the perception of emotions; they do not correctly read the body, vocal and facial expressions of others. With regard to the expression of emotions, it has been shown that they are impaired in expressing emotions through face but up until now, little research has been conducted about their ability to express emotions through spoken language.

To better understand emotion production in both voice and language in Huntington's Disease (HD), we tested 115 individuals: 68 patients (HD), 22 participants carrying the mutant HD gene without any motor symptoms (pre-manifest HD), and 25 controls in a single-centre prospective observational follow-up study. Participants were recorded in interviews in which they were asked to recall sad, angry, happy, and neutral stories. Emotion expression through voice and language was investigated by comparing the identifiability of emotions expressed by controls, preHD and HD patients in these interviews. To assess separately vocal and linguistic expression of emotions in a blind design, we used machine learning models instead of a human jury performing a forced-choice recognition test. Results from this study showed that patients with HD had difficulty expressing emotions through both voice and language compared to preHD participants and controls, who behaved similarly and above chance. In addition, we did not find any differences in expression of emotions between preHD and healthy controls. We further validated our newly proposed methodology with a human jury on the speech produced by the controls. These results are consistent with the hypothesis that emotional deficits in HD are caused by impaired sensori-motor representations of emotions, in line with embodied cognition theories. This study also shows how machine learning models can be leveraged to assess emotion expression in a blind and reproducible way.

* *Corresponding author*. Service de Neurologie, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94000 Créteil, France.
E-mail address: anne-catherine.bachoud-levi@aphp.fr (A.-C. Bachoud Levi).

[1] Equal contribution.

https://doi.org/10.1016/j.cortex.2022.05.024

1. Introduction

Huntington's disease (HD) is a rare autosomal-dominant neurodegenerative disorder that primarily affects the striatum (Tabrizi et al., 2013; Vonsattel et al., 1985). It is characterised by motor, cognitive, and psychiatric disorders, with a median progressive course leading to death within 35 years from symptom onset (Walker, 2007). While motor disorders have been the subject of most studies, they are relatively well tolerated by the patients' entourage (Ho et al., 2011). Conversely, the disruption of patients' social interactions and communication difficulties are one the major causes of patients' social withdrawal and family break-ups. Despite their major importance, they remain poorly understood and insufficiently quantified (Hamilton, 2003; Hartelius et al., 2010; Ho et al., 2011; Jona et al., 2017). Effective communication requires cognitive and linguistic skills, but also social abilities such as the perception and production of emotions. Deficits in emotion perception are a recognised symptom of patients with Huntington's disease. It impairs their ability to decipher and navigate social situations (Bora et al., 2016; Trinkler et al., 2013). However, less is known about these individuals' ability to express emotions, presumably because assessing emotion production is more technically challenging than emotion perception. In particular, the expression of emotions through spoken language has not been explored (C. Kordsachia et al., 2017), despite its critical role for interpersonal communication. This information became even more crucial during the pandemic with the use of telephone communication as the only mean of interaction. The aim of our study is therefore to assess the ability of patients to produce emotions through spoken language.

There is ample evidence that patients with HD have impairments in the recognition of emotional faces (Bora et al., 2016; Kordsachia et al., 2018), body expressions (de Gelder et al., 2008; Zarotti et al., 2019) and voices (Kordsachia et al., 2017; Speedie et al., 1990; Sprengelmeyer et al., 2006), and these impairments extend to negative and positive emotions (Robotham et al., 2011). These deficits can be detected even in the pre-manifest stage of the disease, before the onset of motor symptoms in carriers of the mutant huntingtin gene (Bora et al., 2016; Johnson et al., 2007). Regarding emotion production, only the impairment of facial emotion production has been established (Hayes et al., 2009; Kordsachia et al., 2018; Trinkler et al., 2013, 2017). Given that there is a physiological congruence between body and facial expression in both healthy participants (Zarotti et al., 2019) and in patients with motor impairments (Lenzoni et al., 2020), it can be assumed that body and facial emotion expression are concomitantly impaired. Even though speech plays a critical role in human interpersonal communication, to the best of our knowledge, no study has yet investigated the expression of emotions through spoken language in HD. See Fig. 1 for schematic overview of what is known concerning emotional processes in HD. The communication of emotions through spoken language is reflected both in the voice and the linguistic content (called language here). These two media depend on different brain structures (Friederici, 2017; Guenther, 2016) and can be altered separately. It has been

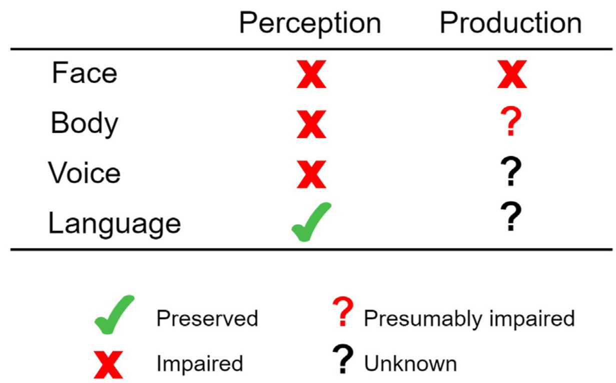

Fig. 1 – Summary of emotional processes in Huntington's disease.

shown that different dimensions of the voice, such as its fundamental frequency, energy, or speech rate, are affected by emotions (Frick, 1985; Paeschke et al., 1999; Scherer, 1995). Similarly, the individual's affective state impacts on linguistic production, including word choice and syntactic structure (Schuller et al., 2011).

Despite their language processing disorders such as syntax or morphology processing (See (Jacquemot & Bachoud-Lévi, 2021) for complete review), it has been shown that individuals with HD do not seem to be impaired in the perception of emotions conveyed by language. Indeed, as observed in a semantic task associating words and emotional content (Hayes et al., 2007) and in the interpretation of a story (Trinkler et al., 2013), participants with HD obtain similar results to those of controls. In contrast, they are impaired in the perception of emotions through voice (Hayes et al., 2007; Kordsachia et al., 2017; Robotham et al., 2011). Assessing the production and the perception of emotions by voice or linguistic content is therefore essential to reach a global vision of emotion processing in HD patients.

Here, we sought to fill these gaps in the literature by testing independently the expression of emotions through voice and language in preHD and HD individuals.

Yet, dissociating emotion expression carried by voice and language in speech represents a methodological and technological challenge. Studies exploring facial emotional expressions (Hayes et al., 2009; Trinkler et al., 2013) used external human scorers to subjectively classify emotions using a forced-choice recognition test. However, this is not applicable to our study since humans cannot perceptually separate the production of emotions by voice and language. Speech can be filtered to make language unintelligible, but at the cost of information loss. Moreover, HD individuals suffer from voice impairments unrelated to emotions that easily distinguish them from controls (Chan et al., 2019; Perez et al., 2018; Riad et al., 2020; Rusz et al., 2014). Indeed, patients exhibit speech disorders that are not emotion-specific, such as dysarthria or speech initiation disorders (Chan et al., 2019; Ludlow et al., 1987; Rusz et al., 2014). Mild speech impairments can even be detected in preHD participants (Chan et al., 2019; Riad et al., 2020; Rusz et al., 2014). Therefore, it is difficult to obtain a blind and dissociated comparison of the emotions conveyed in voice and language with naive listeners. Apart from a human jury, studies on emotional prosody also used statistical analysis of the speech signal (ex: fundamental frequency, energy) (Alhinti et al., 2021; Möbes et al., 2008). However, the link between emotional states and signal characteristics, such as fundamental frequency, intensity, formants are not straightforward (Schuller et al., 2011). This means that we do not know the exact formula that connects emotions, such as anger or joy, to the exact value of the speech signal, especially given the importance of variations across individuals. This limits the information from simple statistical analyses of these signal parameters to study vocal emotion production. The same notion applies to the linguistic information (syntax and semantic) that can be extracted from spoken sentences.

To overcome these difficulties, we developed machine learning emotion recognition models that act as an expert panel to compare the expression of emotions through voice and language between controls, preHD participants and HD participants. This strategy allowed us to assess emotion expression from voice and language separately and without being biased by the disease severity or by spoken language problems unrelated to emotions.

We therefore collected recordings of neutral and emotional speech from healthy controls, preHD individuals and HD individuals. Taking as an assumption that recalling an emotional story elicits the corresponding emotion (Harmon-Jones et al., 2007; Nazareth et al., 2019), we obtained emotional discourses by recording participants while they narrate stories associated with fear, anger and joy. The participants' interviews were segmented and transcribed by speech therapy students. Then, to compare emotions identifiability between controls, preHD individuals and HD individuals, we compared how machine learning models classified emotions from the vocal signal or the annotated text for each group.

2. Material and methods

2.1. Overview

We collected emotional speech at the hospital by asking control, preHD and HD participants, to narrate emotional (anger, sadness, joy) and neutral stories. Interviews were split in stretches and labelled with the name of the elicited emotion. Stretches consist of semantically consistent chunks (Shriberg et al., 2000; Titeux et al., 2021). Stretches from each group (HD, preHD, and controls) and each modality (voice and language) were combined into six sets (3 groups x 2 modalities). We then trained and tested one emotion classifier on each of our six datasets to assess emotion expression through both modalities and for all groups separately (methods are displayed in Fig. 2A). Emotion classifiers are algorithms that can learn how to distinguish emotions on a dataset (either audio signal or text annotation). The classifiers play the role of an expert jury performing a forced-choice recognition test.

We compared the labelling of emotion provided by the classifiers for each stretch to its actual label to measure the accuracy of each classifier. The rationale to use machine learning to examine emotion production is the following: all things being equal in terms of training size, machine learning model and all other parameters, the ability to classify accurately the emotion of a stretch should be equivalent for each group if there is no difference in emotion production. Any difference compared to control will indicate a difference in the group capacity to express emotions.

For the sake of simplicity, we will use the term language for linguistic content in our article. Methods for the voice and language experiments were pre-registered before running any analyses to ensure the validity and avoid inflated results (https://aspredicted.org/OGK_UZA https://aspredicted.org/IQO_YTZ). There were no deviations from the preregistered protocol along the study. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. The link for the open-source code of our analyses is provided at OSF | Vocal and linguistic emotion expression deficits in Huntington's disease.

We further validated our methods with human judgement of the vocal and linguistic production as it is done in classical experiments (Hayes et al., 2009; Trinkler et al., 2013).

2.2. Participants

In this study, 115 participants were included from two observational cohorts (BIOHD NCT01412125 and Repair-HD NCT03119246) at the Hospital Henri-Mondor Créteil, France: 90 participants with at least 36 CAG repeats on the mutant Htt gene (including 22 gene carriers without the manifest disease (preHD), 39 patients at Stage 1, 27 Stage at 2, 2 at Stage 3 according to the total functional score (Shoulson, 1981), and 25 healthy controls (Table 1). Participants carrying the mutant Htt gene were considered as preHD if both their Total Motor Score (TMS) is less than 5 (Tabrizi et al., 2009) and their Total functional capacity (TFC) equals 13 using the Unified Huntington's Disease Rating Scale (UHDRS, Huntington Study Group, 1996). Participants were all French native speakers.

They all signed an informed consent form. Ethical approval was given by the institutional review board from Henri Mondor Hospital (Créteil, France) for the Bio HD study and the CPP Saint Louis French part of the Repair-HD study. It complied with the Helsinki Declaration, current Good Clinical Practice guidelines, and local laws and regulations. None of the participants had any previous or current language, neurological or psychiatric history except HD.

2.3. Clinical evaluation

Concerning the clinical evaluation, participants were assessed by certified examiners through the UHDRS. We reported nine measures classically used for both clinical practice and therapeutic trials (see Table 1): the UHDRS Total Motor Score (TMS), five cognitive assessments (the Symbol Digit Modalities Test (SDMT), the Verbal Fluency test 1-min (VF), and the three components of the Stroop test (word (SW); colour (SC); interference (SI)), and two functional scales (the Total Functional Capacity (TFC) and the Independence scale (UHDRS IS)).

All the demographics and the summary of clinical scores are displayed in Table 1.

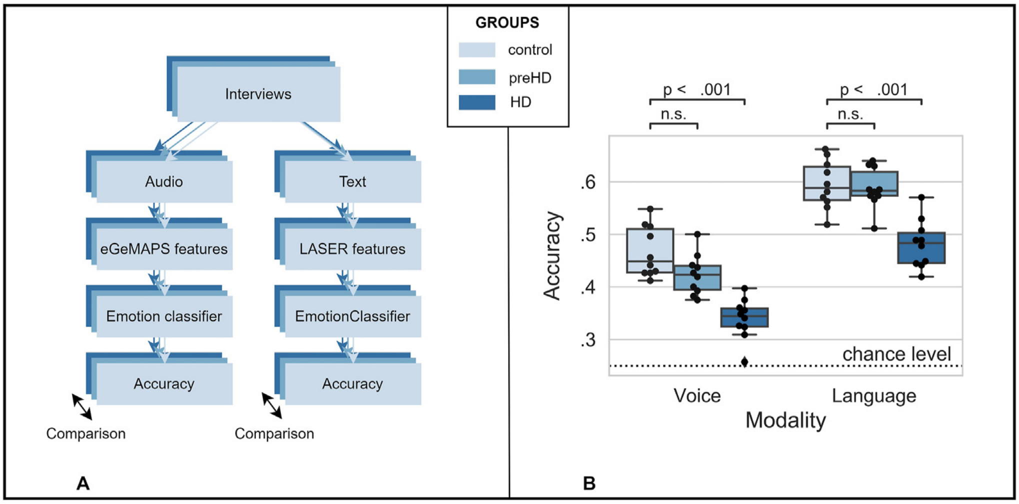

Fig. 2 – Voice and language experiments. A. Method's flowchart. Voice and linguistic content (language) of the interviews are separated to create six datasets (3 groups x 2 modalities). Features are extracted from the data and one emotion classifier is trained and tested on each dataset. Emotion classifiers' accuracy is then compared across groups within each modality. B. Results of accuracy comparison within each modality across the three groups. The HD group accuracy is significantly lower than controls, but this is not the case for the preHD group.

Table 1 – Interviewee demographics and clinical scores.

Sub-groups	Controls	preHD	HD
N	25	22	68
Gender	12F/13M	11F/11M	42F/26M
Age (years)	53.6 (8.8)	49.0 (11.9)	54.3 (10.9)
CAG Triplets	–	41.7 (2.3)	43.8 (3.1)
DBS	–	42 .9 (13.2)	46.1 (13.3)
cUHDRS	–	16.9 (1.4)	9.6 (3.5)
TFC	–	13 (0)	10.5 (2.1)
SW	–	98.8 (12.9)	64.6 (20.7)
SC	–	73.0 (12.1)	46.4 (14.9)
SI	–	43.9 (10.6)	26.0 (10.2)
SDMT	–	51.5 (11.5)	26.5 (9.8)
TMS	–	.5 (1.2)	33.2 (15.8)

DBS: Disease Burden Score; cUHDRS: composite Unified Huntington's Disease Rating Scale; TFC: Total functional capacity; SW: Stroop word reading; SDMT Symbol Digit modality Test; TMS: Total Motor Score; CAG repeats: number of CAG repeats. The Disease Burden Score is computed as age × (CAG-35.5) (Langhben et al., 2007).

2.4. *Data collection*

Interviewers were all neuropsychologists. Participants completed a standardised battery of speech tasks at the hospital. They were asked to narrate both emotional and neutral stories. The stories were prompted by the interviewers asking standardised questions (ex: “May you tell me a story that you find sad, really unhappy, or really depressing”). In most cases, participants spoke for less than 1 min, therefore interviewers were instructed to prompt individuals with pre-defined instructions such as “For instance, you may tell me about an incident heard on the news, a movie, something you saw on television?"). It was clearly stated that they should trigger non personal stories as much as possible. Task 1 (neutral) consisted in describing the latest 24 h and tasks 2, 3, and 4 were designed to elicit emotions by telling a story making the participants sad, angry, and happy, respectively. The speech tasks were separated by non-emotional speech (Automatic recitation of months of the year, the Cookie Theft description, and the storytelling of the Little Red Riding Hood) (McNally et al., 1994) not assessed here. Tasks' order was fixed. All sessions ended with the recall of the happy story, Task 4, to avoid ending the experiment with a negative emotion. The whole session lasted less than 15 min (14.5 ± 6.1 min on average). Patients could interrupt the session anytime.

Participants were recorded in similar acoustic conditions, with a ZOOM H4n Pro recorder, sampled at 44.1 kHz with a 16-bit resolution.

2.5. *Samples preprocessing*

Speech therapists excluded speech samples with too high ambient noise precluding any analysis (two files discarded). Then, they transcribed the language content in text and split by *stretch* the stream of continuous speech from each task in stretches using the software Praat. Annotations were managed with the Seshat platform (Titeux et al., 2021). The annotation of a single interview lasted approximately 8 h, therefore, to ensure the quality of the annotation, we randomly selected five interviews and asked two speech pathologists to annotate them independently. Inter-annotator agreements computed with the Gamma Agreements, denoted γ (Mathet et al., 2015; Titeux & Riad, 2021), for stretch limits and turn-takings were high for the 4 tasks: Neutral ($\gamma = 78.0\% \pm 6.9$), Sad ($\gamma = 84.3\% \pm 7.9$), Angry ($\gamma = 77.6 \pm 4.9$), Happy ($\gamma = 66.0 \pm 19.2$). This allowed to annotate the remaining 110 interviews by a single speech therapist.

Each stretch was labelled with the emotion corresponding to the task it was uttered into (Task 1: neutral, Task 2: sad, Task 3: angry, and Task 4: happy). They were dispatched in data sets according to their group (HD, preHD, and controls) and the modality (voice and language). We balanced the datasets classes to avoid confounding the performance of the machine learning models with the quantity of training data. To do so, stretches were removed randomly from the dataset to ensure the same number of stretches and same number of emotions for the voice and the text. This yielded 1356 stretches for each group, both for the voice and the language modalities.

We then extracted relevant affective information in a fixed-size vector from the audio stretches and linguistic content to be able to automatically classify emotion. We denoted this fixed-size vector the ‘features’.

For the audio stretches, we selected a minimal set of features, the Extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPS) (Eyben et al., 2016) designed by interdisciplinary voice and speech scientists to provide a relevant set of features for affective computing. We chose these features due to their high performance to classify emotions in several voice dataset from different cultures (Eyben et al., 2016). The features are related to energy (e.g., harmonic to noise ratio), to pitch (e.g., fundamental frequency statistics), and articulation (e.g., formant statistics). We extracted these voice features with the openSMILE toolkit (Eyben et al., 2010).

For the text modality, we used the features from the LASER (Language-Agnostic SEntence Representations) sentence embedding model (Artetxe & Schwenk, 2019); LASER being a language-agnostic model which transforms a sentence of words of arbitrary length into a fixed-size vector. It was designed to perform well in a variety of natural language processing tasks in more than 90 languages, especially when the data were scarce to train the models. LASER obtained excellent performance in multi-lingual document classification or natural language inference (find relationships of entailments between sentences).

Then, emotion classifiers were developed for each data set, trained, and tested on the datasets’ features. Emotion classifier algorithms were built with random forests, implemented in scikit-learn (Pedregosa et al., 2011). To train and test emotion classifiers on our datasets we repeated a 10-fold nested cross validation scheme proceeding as follows. We split our data within each dataset in 10 folds. Nine were used for training and then the remaining one as a test set. After 10 permutations between training sets and test sets, we obtained the assessment of the whole data set with a measure of accuracy for each test set, yielding 10 measures of accuracy for each of the six models. The accuracy of an emotion classifier on a set of stimuli (audio or text) labelled with emotions is computed as the percentage of correct predictions (predicted label equals actual label) on this set (see Fig. 2A).

To assess whether the impairment in emotion production through voice in HD is restricted to a specific step during speech production, we ran our audio models with subsets of the features related to energy (Harmonic to noise ratio features, alpha ratio statistic, hammarberg index statistics), pitch (F0 statistics, jitter features, shimmer features), or articulation (MFCC features, formants statistics, spectral flux features). We selected these 3 different sets of features as they represent the main steps to produce fluent articulated speech: respiratory (energy), phonatory (pitch) and articulatory.

2.6. Statistical analysis

To assess the identifiability of emotion expression through voice and language, we compared the accuracy for the two modalities across the three participants ’groups (controls, preHD, and HD). We first used a Kruskal-Wallis test on the three series of accuracies (controls, preHD, and HD), and when it led to significant difference between groups within each modality (rejection of the null hypothesis), we conducted post-hoc Wilcoxon Mann Whitney tests to compare accuracies group by group. The tests were one-tailed, and Bonferroni correction was applied. Because of the structure of the data, traditional tests could not be applied as such, and we resorted to resampling to perform the tests (see Appendix 1 for details). Because the influence of age, gender or task difficulty emotion recognition relies on a complex interaction between these three parameters and the nature of the task, we did not add them in our analyses (Alaerts et al., 2011; Lambrecht et al., 2014; Lenzoni et al., 2020; Richter et al., 2011; Snowden et al., 2008).

2.7. Comparison between machine and human classification

To check whether our method was consistent, 12 healthy scorers with no hearing impairments (31.4 years by mean - s.d. 10.5-, half females, half males) were asked to classify the voice and the text (language) modalities assessed by the machine learning algorithms. They were all French native speakers. For this purpose, annotations were used as the linguistic content and were provided as stretches of text to the scorer. The audio stretches were filtered (we used a Butterworth low pass filter of 4th order, with a cut-off frequency of 250 Hz) to remove intelligibility of words and keep prosody. The 1356 text stretches and the 1356 filtered voice stretches were randomised and divided in six scoring sets for each modality, yielding twelve sets of scoring. Each scorer oversaw one audio scoring set and one text scoring set. Half of them scored first language and then voice (G1), the other half performed the scoring in a reverse order (G2). Text scoring lasted around 15–30 min whereas language scoring for a set lasted between 75 and 90 min for each scorer. We measured agreement between human groups and our model on stretches that were correctly labelled by each individual group. We compared it with agreement between Group 1 (respectively Group 2) on stretches correctly labelled by Group 1 (respectively Group 2). Agreement was measured with the kappa score.

3. Results

Accuracies for each group and each modality are reported in Fig. 2B. We found that our models’ accuracies were significantly lower for HD patients than for controls and preHD for both modalities (corrected p-value < 001 for each statistical test). This shows that HD patients are impaired in expression of emotions through voice and language compared to controls independently from the stage of the disease (results not shown). All types of emotions are impaired across at least one modality (see Fig. 3). Detailed results for each emotion are given in Appendix 2. Accuracies for the preHD group were equivalent to that of controls (voice: corrected p-value = .14, language: corrected p-value >.5). Hence, preHD participants show no sign of impairment in emotion expression through voice and language compared to controls.

The accuracy in classifying emotion was also significantly lower in HD than in controls when restraining the features to the ones tapping energy (see Method), pitch or articulation. In contrast, accuracy in classifying emotion for preHD did not differ from those of controls in any modality (Fig. 4).

The comparison of our method to a human jury highlights its validity (See Fig. 5) The accuracy of the human juries was lower than the machine ones for both modalities (audio: 31.5% on average for human groups *vs* 48% for our model; text: 48% on average for human groups *vs* 60% for our model). We also found that our model and the two human scorers groups correctly classify the same stretches for both modalities (audio: kappa score = .31 on average between human groups (G1 *vs* G2) and .39 on average between each human group and our model; text: .79 on average between the two human groups (G1 *vs* G2) and .7 on average between each human group and our model).

4. Discussion

4.1. *Results summary*

Here, we assessed emotional expression through spoken language of individuals carrying the mutant Htt gene leading to HD in comparison with control participants. As speech combines both the voice (including prosody and vocal signals) and the linguistic content (language), we developed a machine learning method to disentangle emotional expression through voice and through language in a blind and efficient manner. Elicited emotions were classified by emotion classifiers, acting as expert juries. Our models classified stretches of spoken language better than chance for preHD and HD patients, both for voice and language-based models. We found that HD participants have reduced expression of emotions both through voice and language compared to controls and preHD participants. Our machine learning models performed better than human scorers exposed to similar stretches. By studying voice and language separately, we added the missing evidence about the spoken language production of emotion. These results enrich the theories of emotional processing in HD. They also show how machine learning models can be leveraged to study emotion expression.

4.2. *Automated evaluation of emotion expression*

Speech algorithms are currently being developed to detect and classify emotions from spoken language (Akçay & Oğuz, 2020; Picard, 1997; Zhao et al., 2014) and more specifically for individuals affected with neurological disorders (Alhinti et al., 2020; Rusz et al., 2014). They mostly focus on maximizing emotion detection in speech without distinguishing the performance of healthy participants and patients as we did. Yet, these models offer a remarkable opportunity to compare the expressiveness of emotions in spoken language. To the best of our knowledge, machine learning models have never been used to quantify the identifiability of emotions in spoken language of individuals with neurological disorders.

Traditional methods seeking to compare the expression of emotions between patients and controls use either a statistical comparison of pre-defined spoken language markers (Alhinti et al., 2021) or a human jury (Hayes et al., 2009; Trinkler et al., 2013). These methods preclude analysing variation of everyday emotions beyond stereotypical constrained intense emotions in static set-ups (Yitzhak et al., 2020). Our methods allow assessing naturally triggered emotions (Harmon-Jones et al., 2007; Nazareth et al., 2019) without any assumptions between the speech markers and the nature of emotion. It presents three advantages compared to a

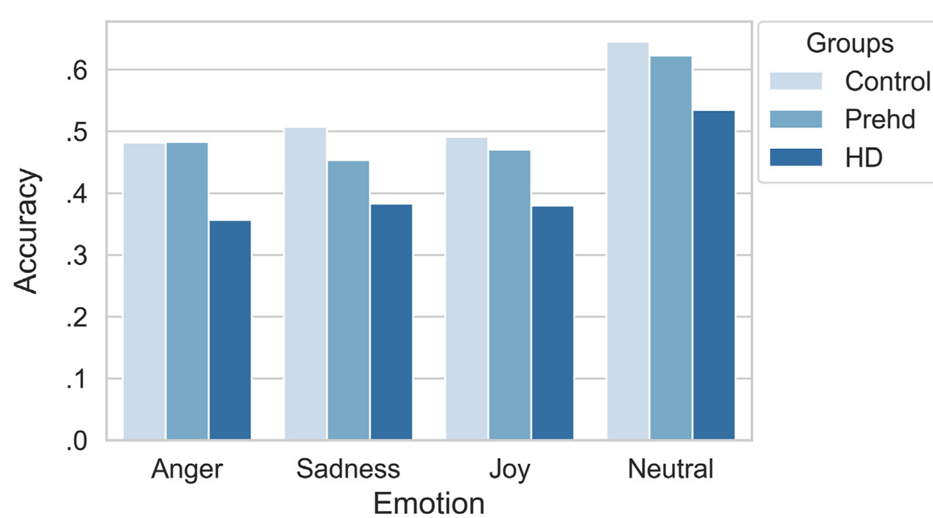

Fig. 3 – Mean of the voice and language accuracy for each emotion and each group. This figure provides a view per emotion, provided through the different splits and the different modalities (Voice and language). Yet, we took the accuracy averaged between the different splits and the different modalities thus precluding a direct statistical comparison but giving a hint on the lowering of all expressions through speech.

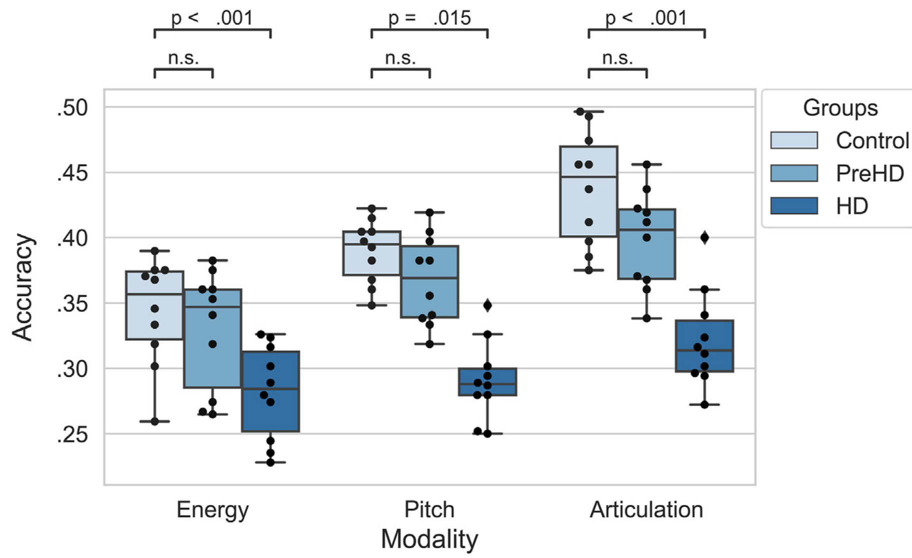

Fig. 4 – Accuracies obtained for each group when restraining the audio features to energy, pitch, or articulation features.

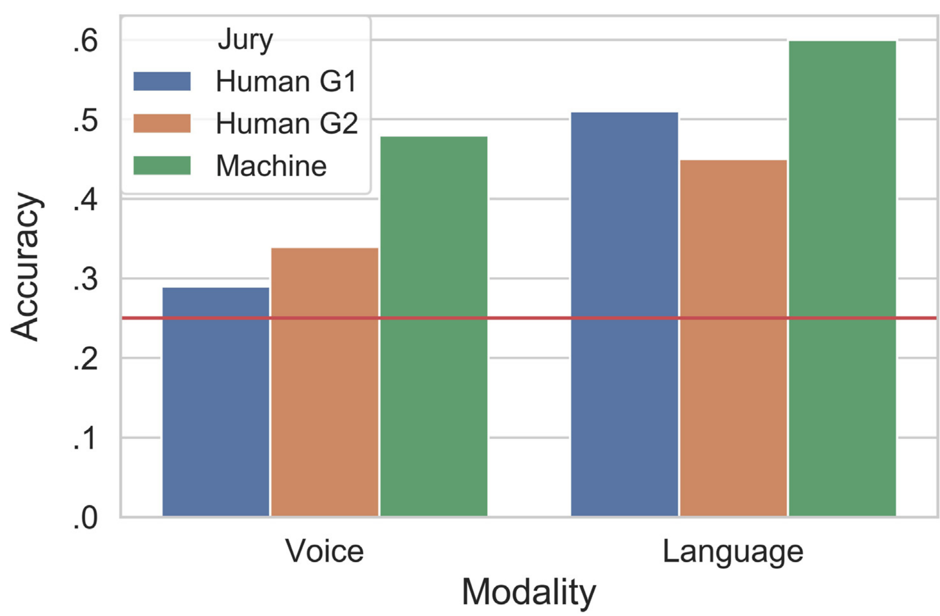

Fig. 5 – Accuracies Human/Machine on a subset of speech segments produced by Healthy Controls.

human jury. First, it is not sensitive to the impact of the disease on the motor aspects of speech, since one model is trained on each population. If the emotions are still expressed despite the motor problems, the models will be able to identify it. This is not the case with a human jury which can be biased by differences across participants unrelated to the expression of emotions. This was particularly sensitive in our study since Patients with Huntington's disease suffer from dysarthria, inappropriate pauses or breathing, and temporal irregularities of speech (Perez et al., 2018; Riad et al., 2020; Rusz et al., 2014) independently of emotional expression. Additionally, machine Learning models allow selective testing of various dimensions that are intertwined in the speech signal, for instance through the choice of features. This enabled us not only to study language and voice separately, but also to compare emotion expression through different speech effectors (energy, pitch, and articulation; see Appendix). Finally, using a standardised method makes it easier to reproduce the analysis, to compare it with new studies; it does not require organising a human experiment with a jury, which is challenging and time consuming. The use of human juries also brings other problems such as huge variability when the sample size is low. Thus, this method could even be adapted to study emotion expression through other modalities such as body and gestures by feeding the model pictures as input to machine learning algorithms.

This shows that machine learning models are great additions to classic approaches using human juries. This opens a room for more straightforward and reproducible research in the study of the production of emotions.

4.3. Embodied cognition & spoken language

Previous studies attempted to explain the reported deficits in expression and perception of emotions in HD with the theoretical framework of embodied cognition (Kordsachia et al., 2017; Trinkler et al., 2017). In the embodied cognition theory, high-level cognitive processes use reactivation of sensory and motor systems. Applied to emotions, it suggests that perceiving others' emotional states involves sensori-motor reexperiencing as opposed to a mere mobilization of abstract conceptual representations of emotions (Niedenthal, 2007). Evidence from studies with neurotypical populations support this view. For instance, the dampening or amplification of facial feedback modulates the accuracy to perceive emotions (Neal & Chartrand, 2011).

The hypothesis made in previous studies is that in HD, the deterioration of motor processing components alters sensorimotor representation of emotions, causing the observed impairments in expression and recognition of emotion. Reported results provide credibility to this argument because they are consistent with three important predictions of this hypothesis.

First, this hypothesis states that expression and recognition mechanisms are caused by the same underlying mechanism. Thus, we shall observe joint deficits in perception and production of emotions, and these deficits should extend to all emotions. Results obtained on expression and recognition of facial emotion in HD are in line with this prediction (Hayes et al., 2009). showed a coupled expression-recognition deficit for facial disgust in HD (Robotham et al., 2011). by equalising the number of positive and negative stimuli, showed that both positive and negative emotions were impaired similarly (Trinkler et al., 2013). and (Trinkler et al., 2017) extended these results and found that facial emotion expression deficits were correlated with impairments in emotion recognition of facial emotions in HD and that these impairments extend to all emotions.

Second, this hypothesis predicts that emotional impairments are tied to motoric impairments rather than dysfunctions in internal experience of emotions or cognitive disorders. Results reported in the literature also support this second prediction. HD patients scored as controls on the alexithymia test (Trinkler et al., 2017) suggesting an intact internal experience of emotions. Additionally, evidence for an intact conceptual understanding of emotions in HD was reported in several studies (Kordsachia et al., 2017). Furthermore (Yitzhak et al., 2020), showed that amongst 4 factors (cognitive screening, motor symptoms, depressive symptoms, and the estimated progression of HD pathology) only motor symptoms were correlated with HD patients' performances in a facial expression recognition task.

Third, following this hypothesis, patients suffering from other movement disorders should exhibit joint deficits in expression and recognition of emotions. Accordingly, it has been shown that in Parkinson's disease (Mermillod et al., 2011), in multiple sclerosis (Henry et al., 2009; Pöttgen et al., 2013) or in myotonic dystrophy (Lenzoni et al., 2020) similar expression/recognition deficits are observed. Consistently, they extend to several emotions and modalities.

Our results on emotion expression through voice provide additional evidence to support the hypothesis of altered sensori-motor representations of emotions in HD. The hypothesis predicts that 1) expression of emotions through motor functions is impaired in HD, and thus expression of emotions through voice should be impaired 2) expression and recognition deficits are joint impairments, and as recognition of emotions through voice is impaired in HD, expression of emotions through voice should also be impaired. Consistently, we found that expression of emotions through voice is impaired in HD. Additionally, we found that this impairment does not seem to be effector (energy, pitch, and articulation) specific.

More difficult to integrate is the preserved perception of emotional language reported in previous studies (Hayes et al., 2007; Trinkler et al., 2013). Following our finding that the expression of emotions through language is impaired, it derogates from the joint perception/production impairment predicted by embodied cognition theory. Nevertheless, as (Winkielman et al., 2018) points out, when the tasks do not require emotional implications, the conceptual channel could remain functional without simulating emotions. HD patients could be impaired in the perception of emotional language but could compensate for their deficit with intact conceptual abilities (Kordsachia et al., 2017). The two hypotheses are not exclusive: although embodied representations and conceptual representations rely on two different networks, their activation is not competitive but complementary. They depend on the context and the purpose of the task (Winkielman et al., 2018). For example, when asked to list properties associated with emotional concepts (e.g., frustration), participants' facial muscles are activated more in an emotional context (they are asked to respond as they would to a good friend) than in a formal context (they are asked to respond as they would to a supervisor) (Niedenthal et al., 2009; Winkielman et al., 2018). Coupled with the interaction between conceptual or embodied language and emotions, this could explain the differences with voice processing.

Another possibility is that the incongruence between perception and production of emotional language is not dependent on emotional processing. The tasks used in perception consisted of classifying either emotional words or emotional stories (Hayes et al., 2007; Trinkler et al., 2013). Although in some cases emotional words may induce emotion, the context of the tasks may not activate emotions. Therefore, it may not be comparable to our production tasks based on the elicitation of emotions with emotional stories. Alternatively, the emotional language impairment could be based on production deficits (decreased fluency, syntactic and semantic deficits in Huntington's disease patients), which are more marked in production than in perception (Ludlow et al., 1987). The production deficit would therefore be independent of the emotional processing deficit.

The processing of action words has been the subject of a similar debate. Patients with Parkinson's and Huntington's disease have difficulties with syntax, verb perception and production, and the coupling between motor and action (Birba et al., 2017). These results are consistent with both the theory of embodied cognition and what (Mahon & Caramazza, 2008) have called the disembodied cognition hypothesis. Counter-examples such as the retention of nouns but the impossible use of tools in apraxic patients or abstract words such as "freedom" that do not induce either simulation or action would confine the embodiment theory to emotion or action words and could not be applied to language in general. This led to an intermediate theory of embodied cognition, the "grounding through interaction", in which conceptual representations can be maintained without motor activations, but sensory and motor systems complement and enrich abstract and symbolic representations (Mahon & Caramazza, 2008). The instantiation of a concept would involve the retrieval of specific sensory and motor information. The "suppression" of sensory and motor systems (as in the case of brain damage) would result in impoverished or "isolated" concepts. From this point of view, sensory and motor information contribute to the "complete" representation of a concept. The activation of sensory and motor processes during conceptual processing is not necessarily "incidental" or "irrelevant" to conceptual processing. The activation of specific sensory and motor representations complements the generality and flexibility of "abstract" and "symbolic" conceptual representations. However, this theory was dedicated to action words and not to emotion words and the entanglement between emotion theories and language processing remains to be clarified.

As our methodology tested voice and language in an equivalent way based on elicited emotions, the impairment of emotion production by language and voice presumably reflects a deficit of emotion expression in HD patients. Previous studies used the framework of embodied cognition to explain the consequences of motor disorders on emotional processes (Trinkler et al., 2017), action language (Birba et al., 2017) or other processes such as mental rotation (Cona et al., 2020), and memory (Dijkstra et al., 2007). Our results are consistent with this framework.

4.4. *Clinical perspectives*

HD patients have difficulties in expressing their emotions through both voice and language may partly account for the disrupted communication between patients and their caregivers (Ho et al., 2011). This difficulty in sharing emotions might create misunderstanding and frustration. This contradicts our previous view that expression of emotions through spoken language was preserved (Trinkler et al., 2017) and could compensate impaired motor expression of emotions. Our results suggest that this might not be the case. Awareness of HD patients' emotion expression impairments should be raised amongst patients and caregivers to improve their interactions.

Second, our results could also be of use for the design of a smart device to monitor Huntington's disease at home. Currently, the disease's follow-up requires regular consultations at the hospital with several neurological, psychiatric, and cognitive tests, which implies both a financial and human cost. Even though HD patients are impaired in emotion expression through spoken language, it is possible to classify emotions in their speech above chance. Thus, an automatic emotion classifier for HD patients could help track their mood and alert caregivers when psychiatric syndromes such as depression, irritability or apathy start developing. They can be hard to spot on a day-to-day basis, and patients and caregivers would benefit from early medical care for these symptoms. In addition, methods are being developed to derive patients' clinical scores from speech (Perez et al., 2018; Riad et al., 2020; Romana et al., 2020). Eventually, these methods could be used to monitor HD at home through the evolution of clinical scores. The identifiability of emotions for each HD individual may constitute an interesting clinical endpoint to monitor as well. It may also potentially be included in clinical trials to spare the capabilities of HD patients' social interactions of HD's individuals rather than on, and not only attenuate motor symptoms.

4.5. *Limitations*

Our study presents some limitations that could be overcome in future works. While being in range with the literature, the number of participants and the amount of data for each participant was limited in our study because of the complexity of retrieving speech data in controlled experiments. Hence, we could not conduct fine grained analysis such as correlations between the identifiability of emotions and motor and cognitive score for each individual.

5. Conclusions

Machine learning allowed us to disentangle voice and language in overt speech in HD. The impact of HD on emotion perception can be potentially translated in models of striatal lesions and tested empirically. This represents an important line of subsequent work, for the specifications of the embodied cognition theory in a more quantitative manner, with direct implications for a complete understanding of emotions in HD. Models might apply to other conditions and emotions expressed by bodies or faces, for example.

Data accessibility

The conditions of our ethics approval, including the ethical consent by participants, do not permit public archiving of anonymized study data. Readers seeking access to the data should contact the corresponding author. Access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data, including a research partnership and the completion of a data transfer agreement provided by the APHP. Legal copyright restrictions prevent public archiving of the UHDRS, which can be obtained from UHDRS® | - Huntington Study Group.

Author contribution

Conceptualization (ACBL, RR, CG, ED); Data curation (CG, RR; XNG); Formal analysis (CG, RR); Funding acquisition (ACBL); Investigation (ACBL; KY, JM; LL; JHB; AS); Methodology (ACBL, ED, RR, CG); Project administration (ACBL); Software (HT, XNG, RR, CG); Supervision (ACBL, ED); Validation (ACBL, ED, RR, CG); Roles/Writing - original draft Writing(ACBL, RR, CG, ED); - review & editing (All).

Acknowledgements

We are very thankful to the participants that took part in our study. We also thank the speech pathologists for the annotations of the speech data. In addition, we thank Laurent Cléret de Langavant, Nicolas Fraisse, Tiffany Monnier, Amin Gharbi, Graça Morgado for their help to carry out this project and helpful discussions to improve this work. Finally we thank Lauren Spring for proofreading. This work is funded in part by the Agence Nationale pour la Recherche (ANR-17-EURE-0017 Frontcog, ANR-10-IDEX-0001-02 PSL*, ANR-19-P3IA-0001 PRAIRIE 3IA Institute) and Grants from Neuratris (ANR-11-INBS-0011), from Facebook AI Research (Research Gift), Google (Faculty Research Award), Microsoft Research (Azure Credits and Grant), and Amazon Web Service (AWS Research Credits).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2022.05.024.

REFERENCES

Akçay, M. B., & Oğuz, K. (2020). Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers. *Speech Communication, 116*, 56–76. https://doi.org/10.1016/j.specom.2019.12.001

Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., & Wenderoth, N. (2011). Action and emotion recognition from point light displays: An investigation of gender differences. *Plos One, 6*(6), Article e20989. https://doi.org/10.1371/journal.pone.0020989

Alhinti, L., Christensen, H., & Cunningham, S. (2021). Acoustic differences in emotional speech of people with dysarthria. *Speech Communication, 126*, 44–60. https://doi.org/10.1016/j.specom.2020.11.005

Alhinti, L., Cunningham, S., & Christensen, H. (2020). Recognising emotions in dysarthric speech using typical speech data. *Interspeech, 2020*, 4821–4825. https://doi.org/10.21437/Interspeech.2020-1825

Artetxe, M., & Schwenk, H. (2019). *Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond*. ArXiv:1812.10464 [Cs] http://arxiv.org/abs/1812.10464.

Birba, A., García-Cordero, I., Kozono, G., Legaz, A., Ibáñez, A., Sedeño, L., & García, A. M. (2017). Losing ground: Frontostriatal atrophy disrupts language embodiment in Parkinson's and Huntington's disease. *Neuroscience and Biobehavioral Reviews, 80*, 673–687. https://doi.org/10.1016/j.neubiorev.2017.07.011

Bora, E., Velakoulis, D., & Walterfang, M. (2016). Social cognition in Huntington's disease: A meta-analysis. *Behavioural Brain Research SreeTestContent1, 297*, 131–140. https://doi.org/10.1016/j.bbr.2015.10.001

Chan, J. C. S., Stout, J. C., & Vogel, A. P. (2019). Speech in prodromal and symptomatic Huntington's disease as a model of measuring onset and progression in dominantly inherited neurodegenerative diseases. *Neuroscience and Biobehavioral Reviews, 107*, 450–460. https://doi.org/10.1016/j.neubiorev.2019.08.009

Cona, G., Casagrande, A., Lenzoni, S., Pegoraro, E., Bozzoni, V., Bello, L., Sorarù, G., Botta, A., & Semenza, C. (2020). The role of motor system in mental rotation: New insights from myotonic dystrophy type 1. *Journal of the International Neuropsychological Society: JINS, 26*(5), 492–502. https://doi.org/10.1017/S1355617719001383

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition, 102*(1), 139–149. https://doi.org/10.1016/j.cognition.2005.12.009

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., Andre, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2016). The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing, 7*(2), 190–202. https://doi.org/10.1109/TAFFC.2015.2457417

Eyben, F., Wöllmer, M., & Schuller, B. (2010). *OpenSMILE – The munich versatile and fast open-source audio feature extractor* (pp. 1459–1462). https://doi.org/10.1145/1873951.1874246

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*(3), 412–429. https://doi.org/10.1037/0033-2909.97.3.412

Friederici, A. D. (2017). *language in our brain: The origins of a uniquely human capacity*. MIT Press.

de Gelder, B., Van den Stock, J., Balaguer, R. de D., & Bachoud-Lévi, A.-C. (2008). Huntington's disease impairs recognition of angry and instrumental body language. *Neuropsychologia, 46*(1), 369–373. https://doi.org/10.1016/j.neuropsychologia.2007.10.015

Guenther, F. H. (2016). *Neural control of speech*. MIT Press.

Hamilton, J. M. (2003). Behavioural abnormalities contribute to functional decline in Huntington's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 74*(1), 120–122. https://doi.org/10.1136/jnnp.74.1.120

Harmon-Jones, E., Amodio, D., & Zinner, L. (2007). *Social psychological methods of emotion elicitation. Handbook of emotion elicitation and assessment*.

Hartelius, L., Jonsson, M., Rickeberg, A., & Laakso, K. (2010). Communication and Huntington's disease: Qualitative interviews and focus groups with persons with Huntington's disease, family members, and carers. *International Journal of Language & Communication Disorders, 45*(3), 381–393. https://doi.org/10.3109/13682820903105145

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2007). Disgust and Huntington's disease. *Neuropsychologia, 45*(6), 1135–1151. https://doi.org/10.1016/j.neuropsychologia.2006.10.015

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2009). Production of spontaneous and posed facial expressions in patients with Huntington's disease: Impaired communication of disgust. *Cognition & Emotion, 23*(1), 118–134. https://doi.org/10.1080/02699930801949090

Henry, J. D., Phillips, L. H., Beatty, W. W., Mcdonald, S., Longley, W. A., Joscelyne, A., & Rendell, P. G. (2009). Evidence for deficits in facial affect recognition and theory of mind in multiple sclerosis. *Journal of the International Neuropsychological Society, 15*(2), 277–285. https://doi.org/10.1017/S1355617709090195

Ho, A. K., Hocaoglu, M. B., & European Huntington's Disease Network Quality of Life Working Group. (2011). Impact of Huntington's across the entire disease spectrum: The phases and stages of disease from the patient perspective. *Clinical Genetics, 80*(3), 235–239. https://doi.org/10.1111/j.1399-0004.2011.01748.x

Huntington Study Group. (1996). Unified Huntington's Disease Rating Scale: reliability and consistency. *Huntington Study Group. Movement Disorders. 1996, 11*, 136–142. https://doi.org/10.1002/mds.870110204

Jacquemot, C., & Bachoud-Lévi, A.-C. (2021). Striatum and language processing: Where do we stand? *Cognition, 213*, Article 104785. https://doi.org/10.1016/j.cognition.2021.104785

Johnson, S. A., Stout, J. C., Solomon, A. C., Langbehn, D. R., Aylward, E. H., Cruce, C. B., Ross, C. A., Nance, M., Kayson, E., Julian-Baros, E., Hayden, M. R., Kieburtz, K., Guttman, M., Oakes, D., Shoulson, I., Beglinger, L., Duff, K., Penziner, E., Paulsen, J. S., & Predict-HD Investigators of the Huntington Study Group. (2007). Beyond disgust: Impaired recognition of negative emotions prior to diagnosis in Huntington's disease. *Brain: a Journal of Neurology, 130*(Pt 7), 1732–1744. https://doi.org/10.1093/brain/awm107

Jona, C. M. H., Labuschagne, I., Mercieca, E.-C., Fisher, F., Gluyas, C., Stout, J. C., & Andrews, S. C. (2017). Families affected by Huntington's disease report difficulties in communication, emotional involvement, and problem solving. *Journal of Huntington's Disease, 6*(3), 169–177. https://doi.org/10.3233/JHD-170250

Kordsachia, C. C., Labuschagne, I., Andrews, S. C., & Stout, J. C. (2018). Diminished facial EMG responses to disgusting scenes and happy and fearful faces in Huntington's disease. *Cortex, 106*, 185–199. https://doi.org/10.1016/j.cortex.2018.05.019

Kordsachia, C., Labuschagne, I., & Stout, J. (2017 Feb). Beyond emotion recognition deficits: A theory guided analysis of emotion processing in Huntington's disease. *Neuroscience & Biobehavioral Reviews, 73*, 276–292. https://doi.org/10.1016/j.neubiorev.2016.11.020. Epub 2016 Nov 29.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition & Emotion, 28*(3), 452–469. https://doi.org/10.1080/02699931.2013.837378

Langbehn, D. R., Paulsen, J. S., & Huntington Study Group. (2007). Predictors of diagnosis in Huntington disease. *Neurology, 68*, 1710–1717. https://doi.org/10.1212/01.wnl.0000261918.90053.96

Lenzoni, S., Bozzoni, V., Burgio, F., de Gelder, B., Wennberg, A., Botta, A., Pegoraro, E., & Semenza, C. (2020). Recognition of emotions conveyed by facial expression and body postures in myotonic dystrophy (DM). *Cortex, 127*, 58–66. https://doi.org/10.1016/j.cortex.2020.02.005

Ludlow, C. L., Connor, N. P., & Bassich, C. J. (1987). Speech timing in Parkinson's and Huntington's disease. *Brain and Language, 32*(2), 195–214. https://doi.org/10.1016/0093-934X(87)90124-6

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1–3), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Mathet, Y., Widlöcher, A., & Métivier, J.-P. (2015). *The unified and holistic method Gamma (γ) for inter-annotator agreement measure and alignment*. Computational Linguistics. https://doi.org/10.1162/COLI_a_00227

McNally, R., Litz, B., Prassas, A., Shin, L., & Weathers, F. (1994). Emotional priming of autobiographical memory in post-traumatic stress disorder. *Cognition & Emotion - COGNITION EMOTION, 8*, 351–367. https://doi.org/10.1080/02699939408408946

Mermillod, M., Vermeulen, N., Droit-Volet, S., Jalenques, I., Durif, F., & Niedenthal, P. (2011). Embodying emotional disorders: New hypotheses about possible emotional consequences of motor disorders in Parkinson's disease and tourette's syndrome. *ISRN Neurology, 2011*, 1–6. https://doi.org/10.5402/2011/306918

Möbes, J., Joppich, G., Stiebritz, F., Dengler, R., & Schröder, C. (2008). Emotional speech in Parkinson's disease. *Movement Disorders : Official Journal of the Movement Disorder Society, 23*, 824–829. https://doi.org/10.1002/mds.21940

Nazareth, D., Tournier, E., Leimkötter, S., Janse, E., Heylen, D., Westerhof, G., & Truong, K. (2019). *An acoustic and lexical analysis of emotional valence in spontaneous speech: Autobiographical memory recall in older adults* (pp. 3287–3291). https://doi.org/10.21437/Interspeech.2019-1823

Neal, D. T., & Chartrand, T. L. (2011). Embodied emotion perception: Amplifying and dampening facial feedback modulates emotion perception accuracy. *Social Psychological and Personality Science, 2*(6), 673–678. https://doi.org/10.1177/1948550611406138

Niedenthal, P. M. (2007). Embodying emotion. *Science (New York, N.Y.), 316*(5827), 1002–1005. https://doi.org/10.1126/science.1136930

Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology, 96*(6), 1120–1136. https://doi.org/10.1037/a0015574

Paeschke, A., Kienast, M., & Sendlmeier, W. F. (1999). *F0-CONTOURS in emotional speech* (Vol. 4).

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Louppe, G., Prettenhofer, P., Weiss, R., Weiss, R. J., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*.

Perez, M., Jin, W., Le, D., Carlozzi, N., Dayalu, P., Roberts, A., & Provost, E. M. (2018). Classification of Huntington disease using acoustic and lexical features. *Interspeech, 2018*, 1898–1902. https://doi.org/10.21437/interspeech.2018-2029

Picard, R. W. (1997). *Affective computing*. MIT Press.

Pöttgen, J., Dziobek, I., Reh, S., Heesen, C., & Gold, S. M. (2013). Impaired social cognition in multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry, 84*(5), 523–528. https://doi.org/10.1136/jnnp-2012-304157

Riad, R., Titeux, H., Lemoine, L., Montillot, J., Bagnou, J. H., Cao, X. N., Dupoux, E., & Bachoud-Lévi, A.-C. (2020). *Vocal markers from sustained phonation in Huntington's Disease*. ArXiv:2006.05365 [Cs, Eess] http://arxiv.org/abs/2006.05365.

Richter, D., Dietzel, C., & Kunzmann, U. (2011). Age differences in emotion recognition: The task matters. *The Journals of Gerontology: Series B, 66B*(1), 48–55. https://doi.org/10.1093/geronb/gbq068

Robotham, L., Sauter, D. A., Bachoud-Lévi, A.-C., & Trinkler, I. (2011). The impairment of emotion recognition in Huntington's disease extends to positive emotions. *Cortex, 47*(7), 880–884. https://doi.org/10.1016/j.cortex.2011.02.014

Romana, A., Bandon, J., Carlozzi, N., Roberts, A., & Provost, E. M. (2020). Classification of manifest Huntington disease using vowel distortion measures. *Interspeech, 2020*, 4966–4970. https://doi.org/10.21437/interspeech.2020-2724

Rusz, J., Klempíř, J., Tykalová, T., Baborová, E., Čmejla, R., Růžička, E., & Roth, J. (2014). Characteristics and occurrence of speech impairment in Huntington's disease: Possible influence of antipsychotic medication. *Journal of Neural Transmission (Vienna, Austria: 1996), 121*(12), 1529–1539. https://doi.org/10.1007/s00702-014-1229-8

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235–248. https://doi.org/10.1016/S0892-1997(05)80231-0

Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication, 53*(9), 1062–1087. https://doi.org/10.1016/j.specom.2011.01.011

Shriberg, E., Stolcke, A., Hakkani-Tür, D., & Tür, G. (2000). Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication, 32*(1), 127–154. https://doi.org/10.1016/S0167-6393(00)00028-5

Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia.

Neuropsychologia, 46(11), 2638–2649. https://doi.org/10.1016/j.neuropsychologia.2008.04.018

Speedie, L. J., Brake, N., Folstein, S. E., Bowers, D., & Heilman, K. M. (1990). Comprehension of prosody in Huntington's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 53*(7), 607–610. https://doi.org/10.1136/jnnp.53.7.607

Sprengelmeyer, R., Schroeder, U., Young, A. W., & Epplen, J. T. (2006). Disgust in pre-clinical Huntington's disease: A longitudinal study. *Neuropsychologia, 44*(4), 518–533. https://doi.org/10.1016/j.neuropsychologia.2005.07.003

Tabrizi, S. J., Langbehn, D. R., Leavitt, B. R., Roos, R. A., Durr, A., Craufurd, D., Kennard, C., Hicks, S. L., Fox, N. C., Scahill, R. I., Borowsky, B., Tobin, A. J., Rosas, H. D., Johnson, H., Reilmann, R., Landwehrmeyer, B., & Stout, J. C. (2009). Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: Cross-sectional analysis of baseline data. *Lancet Neurology, 8*(9), 791–801. https://doi.org/10.1016/S1474-4422(09)70170-X

Tabrizi, S. J., Scahill, R. I., Owen, G., Durr, A., Leavitt, B. R., Roos, R. A., Borowsky, B., Landwehrmeyer, B., Frost, C., Johnson, H., Craufurd, D., Reilmann, R., Stout, J. C., & Langbehn, D. R. (2013). Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: Analysis of 36-month observational data. *Lancet Neurology, 12*(7), 637–649. https://doi.org/10.1016/S1474-4422(13)70088-7

Titeux, H., & Riad, R. (2021). pygamma-agreement: Gamma γ measure for inter/intra-annotator agreement in Python. *The Journal of Open Source Software, 6*, 2989. https://doi.org/10.21105/joss.02989

Titeux, H., Riad, R., Cao, X.-N., Hamilakis, N., Madden, K., Cristia, A., Bachoud-Lévi, A.-C., & Dupoux, E. (2021). *Seshat: A tool for managing and verifying annotation campaigns of audio data.* ArXiv:2003.01472 [Cs] http://arxiv.org/abs/2003.01472.

Trinkler, I., Cleret de Langavant, L., & Bachoud-Lévi, A.-C. (2013). Joint recognition-expression impairment of facial emotions in Huntington's disease despite intact understanding of feelings. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 49*(2), 549–558. https://doi.org/10.1016/j.cortex.2011.12.003

Trinkler, I., Devignevielle, S., Achaibou, A., Ligneul, R. V., Brugières, P., Cleret de Langavant, L., De Gelder, B., Scahill, R., Schwartz, S., & Bachoud-Lévi, A.-C. (2017). Embodied emotion impairment in Huntington's Disease. *Cortex, 92*, 44–56. https://doi.org/10.1016/j.cortex.2017.02.019

Vonsattel, J. P., Myers, R. H., Stevens, T. J., Ferrante, R. J., Bird, E. D., & Richardson, E. P. (1985). Neuropathological classification of Huntington's disease. *Journal of Neuropathology and Experimental Neurology, 44*(6), 559–577. https://doi.org/10.1097/00005072-198511000-00003

Walker, F. O. (2007). Huntington's disease. *The Lancet, 369*(9557), 218–228. https://doi.org/10.1016/S0140-6736(07)60111-1

Winkielman, P., Coulson, S., & Niedenthal, P. (2018). Dynamic grounding of emotion concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1752), Article 20170127. https://doi.org/10.1098/rstb.2017.0127

Yitzhak, N., Gurevich, T., Inbar, N., Lecker, M., Atias, D., Avramovich, H., & Aviezer, H. (2020). Recognition of emotion from subtle and non-stereotypical dynamic facial expressions in Huntington's disease. *Cortex, 126*, 343–354. https://doi.org/10.1016/j.cortex.2020.01.019

Zarotti, N., Fletcher, I., & Simpson, J. (2019). New perspectives on emotional processing in people with symptomatic Huntington's disease: Impaired emotion regulation and recognition of emotional body language. *Archives of Clinical Neuropsychology, 34*(5), 610–624. https://doi.org/10.1093/arclin/acy085

Zhao, S., Rudzicz, F., Carvalho, L. G., Marquez-Chin, C., & Livingstone, S. (2014). Automatic detection of expressed emotion in Parkinson's Disease. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4813–4817). https://doi.org/10.1109/ICASSP.2014.6854516