INSTITUTO
UNIVERSITÁRIO
DE LISBOA

The time course of authenticity and valence perception in nonverbal emotional vocalizations

Elisabete Sofia Nabais de Oliveira de Freitas e Menezes

Supervisor:

César Lima, PhD, Assistant Professor
ISCTE-IU

November, 2022

iscte

CIÊNCIAS SOCIAIS
E HUMANAS

Department of Social and Organizational Psychology

The time course of authenticity and valence perception in nonverbal emotional vocalizations

Elisabete Sofia Nabais de Oliveira de Freitas e Menezes

Supervisor:

César Lima, PhD, Assistant Professor
ISCTE-IUL

November, 2022

To my granddaughters Laura and Leonor

May life always enable you to identify authentic smiles above all.

Acknowledgements

First of all, I would like to express my endless gratitude to my supervisor César Lima for having accepted the tricky task of supervising me and, since that moment, for his permanent support, guidance and patience. I would also like to thank, from the bottom of my heart, my *Emotions' Guardian Angels* Matilde Rebello de Andrade, Aíssa Baldé and Ana Isabel Correia for their commitment, expertise and friendship along this path. Without them I would have never been able to conclude this dissertation. Thank you Pedro Almeida and António Álvaro Soares for the time, the discussions, the works we had together and that motivated me to go further on the road of Emotions. Also, a huge Thank You to all participants that took part in the study and enabled this research to be possible.

Finally, the most special thanks go to my family: Paulo Menezes, Maria Helena Oliveira, André Menezes, Diogo Menezes, Chris Rademacher and Ana Oliveira, who never stop believing me and my craziest dreams even when the world thinks they will not come true. Thanks to their love and comforting smiles my dreams have always come true!

Abstract

There is evidence that the recognition of sadness and happiness in nonverbal vocalizations reaches an adult standard before the recognition of anger and fear, and that men and women are equally good at recognizing emotions, regardless of whether male or female speakers produce them. Still, there is no evidence regarding how much time we need to identify the authenticity of vocal emotional expressions, as well as the type of vocalization itself. How much acoustic information do we need to perceive if a vocal expression, such as a laughter, is authentic or voluntary? How long does it take to perceive if its laughter or crying? The present study addresses these questions. The main objective is to determine the time course of authenticity and type of vocalization recognition in laughter and crying sounds. For this purpose, the procedure was done using a gating paradigm and a sample of 395 participants. Results showed that the recognition accuracy of nonverbal vocalizations improves with the increase of the gate duration, and that the identification of the type of vocalization (laughter vs. crying) happens at earlier stages than the identification of their authenticity (authentic vs. voluntary).

Keywords: nonverbal vocalization, authenticity, valence, time course

Classification APA PsycINFO:

2300 Human Experimental Psychology

2346 Attention

2360 Motivation & Emotion

Resumo

Há evidências de que o reconhecimento de tristeza e felicidade em vocalizações não verbais atinge um padrão adulto antes do reconhecimento de raiva e medo, e que homens e mulheres são igualmente bons a reconhecer emoções, independentemente de estas emoções serem produzidas por falantes do sexo masculino ou do sexo feminino. Ainda assim, não há evidências de quanto tempo precisamos para identificar a autenticidade das expressões emocionais vocais, bem como o tipo de vocalização em si. De quanta informação acústica precisamos para perceber se uma expressão vocal, tal como um riso ou choro, é autêntica ou voluntária? Quanto tempo se demora a perceber se é riso ou choro? O presente estudo aborda estas questões. O objetivo principal é determinar o tempo de reconhecimento da autenticidade e tipo de vocalização em estímulos de riso e choro. Para tal, foi utilizado um paradigma de *gating* e uma amostra de 395 participantes. Os resultados mostraram que a precisão do reconhecimento de vocalizações não verbais melhora com o aumento da duração do *gate* e que a identificação do tipo de vocalização (riso vs. choro) ocorre em fases mais precoces do que a identificação da sua autenticidade (autêntica vs. voluntária).

Palavras-chave: vocalizações não verbais, autenticidade, valência, espaço temporal

Classificação APA PsycINFO:

2300 Psicologia Experimental Humana

2346 Atenção

2360 Motivação e Emoção

Contents

Table of Figures

Tables

1. Introduction

Human beings are social animals who communicate via verbal and nonverbal information. Language is part of that information and it can be defined as a system used to communicate both cognitively and emotionally using sounds or gestures, and it depends on both cortical and subcortical brain networks (Johnson, 2017). Language has a crucial role in communication, whether by means of teaching or by learning the best ways to make one's message clearer. Nevertheless, the study of nonverbal information has been conquering its place in the field of research in communication over the last decades. Facial expressions, body postures, touch, and voice cues make nonverbal communication an appealing ground for research (Belin & Gosselin, 2008; Ekman and Keltner, 1997; Lima, Castro & Scott, 2013; Nummenmaa & Calvo, 2015; Sauter, 2017).

1.1 The Human Voice

The voice is a vital instrument for transmission of verbal and nonverbal information, and it is considered a powerful tool for social interactions (Sauter, 2010) and the carrier of speech (Belin, 2004). It provides crucial information regarding the emotional states of the speaker (Neves, Cordeiro, Scott, Castro, & Lima, 2018; Scherer, 1995). It is also used to identify other socially relevant aspects, such as the identity, sex and age of the speaker (Belin et al., 2004).

The voice is an instrument including several distinct and relevant acoustic elements (intensity, pitch, timing, and timbre), perceived by two of the five senses, because we can hear and feel it, in the sense that sound production is felt through vibrations in the head, neck and chest (Titze, 2000). The vital role of the voice in our lives starts even before birth and it is expressed by each of us when we are born and use the sound produced by vocal folds vibration to communicate and express emotions. The voice produces spoken words that result from three components: resonance, articulation, and vocal sound.

In vocal communication, emotions can be expressed in various ways, including lexico-semantic information in speech (by expressing emotions through words), emotional prosody (emotionally inflected speech) and purely nonverbal vocalizations (e.g., laughter, crying). Emotions like happiness, for example, can be defined in vocal terms by laughter, and the same goes for sadness when expressed by sobbing or crying (Titze, 2000).

1.2 Prosody

Prosody can be defined as the nonverbal aspects of speech (Nordstrom, 2019). It is linked to vocal qualities of speech, and includes intonation, stress, timing and timbre (Thomson, 2006). Authors such as Cowen et al. (2019) have studied how people can recognize different emotions (12 in their study) from speech prosody. Their results showed that different emotions can be preserved across cultures even when they have distinct acoustic correlates.

When speech prosody is used to convey information other than words with the function of communicating emotion (Nordstrom, 2019), it is named emotional prosody, and in research, stimuli can include words/sentences or pseudo-words/sentences (Vasconcelos et al., 2017). Jiang et al. (2015) define emotional prosody as vocal expressions of emotion embedded in human speech. According to Titze (2000), it is a way to communicate vocal emotions via suprasegmental modulations of speech.

1.3 Nonverbal vocalizations

Over the last 75 years, many studies have examined listeners' ability to recognize a speaker's affect state or attitude from verbal voice samples (e.g., numbers, letters of the alphabet, nonsense syllables, or standardized sentences; Scherer, 1995). However, the communication of vocal emotions also happens via purely nonverbal vocalizations.

Nonverbal emotional vocalizations (e.g., laughter, cries, screams, moans, sighs) are brief vocal expressions that do not involve linguistic content, even though they can convey important emotional information (e.g., Sauter & Fischer, 2018). They are highly variable and can reflect distinct socio-emotional processes for different forms of the same expression, e.g., a laugh can be authentic or acted out (Neves et al., 2018). Scott et al. (2009) referred to nonverbal vocalizations as the most phylogenetically continuous of all expressive channels, differing from speech prosody for not being constrained by the articulatory movements required by speech production.

According to some authors, speech prosody (Juslin & Laukka, 2001) and nonverbal vocalizations (Sauter et al. 2010) rely on similar acoustic cues (e.g., duration, fundamental frequency, intensity) to express emotional meaning. On the other hand, Scott et al. (2009) refer that nonverbal vocalizations involve distinct production mechanisms from speech, resulting in distinct acoustic properties.

Most phonemes are voiced, including vowels, semivowels, and nasals (Cowen, Laukka, Elfenbein, et al 2019). These empower language while communicating, as prosody connects with language in the shape of words to transmit emotional feelings and attitudes (Laukka et al., 2016).

Some authors also compare human nonverbal vocalizations to affect vocalizations produced by other animals (Scherer, 1995), and nonverbal vocalizations can vary according to valence (Lima, Alves, Scott, & Castro, 2014), arousal (Sauter et al., 2010), authenticity (genuine vs. non-genuine; McGettigan et al., 2015), and emotional contagion (Neves et al., 2018).

Lima et al. (2014) state that nonverbal vocalizations communicate a range of positive and negative emotions, and Belin et al. (2008) define them as affective bursts, adding that when compared to speech prosody, they are described as a representation of more primitive expressions of emotions and as the auditory counterpart of facial emotions (Belin et al. 2004). Until recently, studies on nonverbal communication of emotion were mainly focused on facial expressions (Ekman, 1997; Lavan, 2015; Nummenmaa & Calvo, 2015). However, other studies and research also suggest good recognition of several positive emotions, such as affection, happiness or interest (Laukka et al., 2013); triumph, pleasure, amusement or relief (Sauter et al. (2010) from vocal expressions such as nonverbal vocalizations (e.g., Cowen et al., 2018; Laukka et al., 2013). Sauter and Scott (2007) showed that when compared to facial expressions, nonverbal vocalizations can convey a larger number of positive emotional states.

However, there are not many studies delineating the acoustic patterns associated with positive emotions (Sauter, 2017). Vasconcelos et al. (2017) focused on the valence of emotions, comparing how effectively positive and negative emotions can be recognized through nonverbal emotional vocalizations. They state that positive emotions (such as laughter, surprise, and pleasure) are more easily recognized through nonverbal vocalizations compared to negative emotions (such as fear, anger, sadness, and pain). On the other hand, Castiajo and Pinheiro (2019) associated sadness and disgust with the highest accuracy concerning the decoding of nonverbal vocalizations, whereas fear, anger and happiness were the least accurately decoded ones.

Nonverbal vocalizations can provide us with information about other people's needs and intentions (Scott et al., 2014), which brings us to the first sentence of this chapter: 'the human being is a social animal', whose social interactions are structured and regulated not only by verbal language but also by nonverbal cues (Keltner & Kring, 1998) that guide the behaviour of individuals and produce adequate responses according to context (Pires, 2019). To illustrate this statement, some studies (Bryant, 2008; Jiang et al., 2015; Laukka et al., 2013) have supported the hypothesis that vocal emotional expressions of basic affective categories manifest themselves in relatively similar ways across quite different cultures. In the present study, we focused on two nonverbal emotional vocalizations – laughter and crying –, which are described in detail in the following paragraphs.

1.4 Laughter

Laughter is a social expression of emotion, typically associated with bonding, agreement, affection, and emotional regulation (Scott et al., 2014), which has been identified in several mammal species in addition to humans (Davila-Ross et al. 2011). In humans, laughter is considered the only positive vocal emotional expression recognized across cultures and in geographically distinct groups (Sauter et al. 2010). According to McGettigan et al. (2015), laughter can work as an essential behaviour to help to ease negative emotional experiences and in situations of significant stress. It can be used as a coping strategy (e.g., nervous laughter) and occurs more often when we are with other people than when we are alone (Scott et al., 2014). The laughter signal can vary in intensity, naturalness, and authenticity, among other variables (McKeown et al., 2015).

Laughter production is characterized by coordinated actions between the respiratory system and the laryngeal and rib muscles (Bryant & Aktipis, 2014) and by very rapid contractions of the intercostal muscles, resulting in large exhalations followed by individual bursts of laughter (McGettigan et al., 2015). Based on studies of pathological laughter, Wild et al. (2003) describe a model of laughter production that identifies a coordinating centre for laughter in the brainstem comprising the periaqueductal grey and the upper reticular formation in the control of changes in facial expressions, respiration, and vocalization. According to this model, the lateral premotor cortices participate in the primary production of laughter and suppress and regulate authentic laughter vocalizations. According to some authors (e.g., Lavan et al., 2017; McGettigan et al., 2015; Scott et al., 2014), there is an increased activation of specific brain areas in terms of the perception of each type of laughter: voluntary laughter generates more activity than authentic laughter in the anterior medial prefrontal cortex due to the presence of social ambiguity in the acoustic information. Authentic laughter, on the other hand, is suggested by some authors (Pisanski et al., 2016; Scott et al., 2014) as having its origin in a complex set of midline structures involved in innate vocalizations (e.g., periaqueductal grey).

As described by McGettigan et al. (2015), laughter can be divided into two types, according to the way they are elicited. It can be either driven by outside events (authentic) or by intentional communicative acts (voluntary). Scott et al. (2014) added that these two different types of laughter can be found in humans and chimpanzees and can be observed in brain imaging studies of laughter. The main differences between authentic and voluntary laughter are higher pitch, longer duration, and visible spectral characteristics in authentic laughter, which are distinct from voluntary laughter, and the latter is more nasal than authentic laughter. Perceptually, authentic laughter is perceived as more ‘real’ than voluntary laughter and as more positive and higher in arousal (Neves et al., 2018). Neves (2018) also proposes that authentic laughter reflects a less controlled and genuinely felt emotion, whereas voluntary laughter transmits more deliberate communicative acts. Other authors also state that voluntary and authentic laughter are perceptually and acoustically distinct (Bryant & Aktipis, 2014).

In a study on the recognition of emotions, Amorim et al. (2021) report higher recognition of happiness accuracy compared to other emotions due to the high social relevance of laughter. Recent studies have been focusing on the ability to detect the authenticity of emotional expressions (Lima et al., 2021), and results show that listeners are able to accurately distinguish between authentic and voluntary laughs (e.g., Bryant et al., 2018; Lavan et al., 2016; Neves et al., 2018). Here, our primary aim will be to find out what is the amount of time that people need to recognize that a laugh is authentic or voluntary.

1.5 Crying

The human being is the only species that produces tears in response to emotional events (Vingerhoets, 2013). Children cry to communicate their survival and developmental needs (Miceli & Castelfranchi, 2003; Sadoff, 1996). Crying involves respiratory, laryngeal, and supralaryngeal articulatory activity (Bylsma et al., 2019) and can be expressed in two ways: through facial signals, like tears, and through vocalizations (Pires, 2019). There are not many studies focusing on crying and its emotional recognition (e.g. Simons et al., 2013; Vingerhoets & Bylsma, 2015). Therefore, much research needs to be carried out to better characterize this social expression of emotion. Still, authors describe crying as a relevant marker regarding emotional events in humans' lives (Simons et al., 2013), whether the events are positive or negative. Crying can be triggered by different types of emotion: sadness, happiness, anger, guilt, humiliation, frustration, or anxiety (Frey, 1985; Miceli & Castelfranchi, 2003). However, some authors describe crying not as an expression of emotions but as a communicative tool used to transmit social motives to others (e.g., Fridlund, 1991)

Similarly to laughter, current research identifies two types of crying: authentic and voluntary (Anikin & Lima, 2018; Costa, 2018; Lavan et al., 2015). Authentic crying is more straightforward to identify in tears (Provine et al., 2009) than in terms of sound. However, it is still complex to obtain authentic crying production to be used in experimental settings and generated via emotion induction, contrarily to what happens with authentic laughter which is relatively easy to obtain in a laboratory (Pinheiro et al, 2021) and usually generates stronger physiological responses (Lima et al, 2021). In this regard, some authors refer to a possible existing ambiguity when sometimes mistaking authentic crying for authentic laughter (Lavan et al., 2015).

Concerning the situations in which the two types of crying can occur, we have already mentioned the different types of emotion that can trigger authentic crying. Regarding voluntary crying, literature mainly focuses on situations of emotional manipulation, such as when children try to achieve their goals by pressuring their parents or other adults in charge (Gračanin et al., 2018; Simons et al., 2013; Vingerhoets & Scheirs, 2000). Compared to laughter and the recognition of positive emotions such as joy or happiness, studies suggest a faster recognition of sadness in crying due to the relevant role of sadness in social behavior learning (Blasi et al., 2015).

1.6 Authenticity

Over the last decade, there has been increasing research on the field of authenticity of emotions (Anikin & Lima, 2017; Costa, 2018; Drolet et al., 2012 and 2014; Lavan et al., 2017; Neves, 2018). Lavan and colleagues (2017) highlight the importance of distinguishing if emotions are spontaneously or voluntarily expressed in everyday social interactions.

Regarding previous research on authenticity, several authors have presented different approaches, some focusing on prosody, others on nonverbal vocalizations. Most studies on authenticity involve actors who portray an emotion with a set of guidelines (voluntary portrayal of emotion) (Costa, 2018). The following study used different techniques to identify authenticity portraying voice recorded stimuli performed by actors. In a study by Scheiner and Fischer (2011), the recordings used were made from real emotions experienced by individuals as they went through specific situations or while talking about real past events. Individuals were also asked to describe their emotional states. By recording these real-life situations, the authors were able to present 80 non-instructed "authentic" speech recordings of four emotional expression categories (20 of anger, 20 of sadness, 20 of joy, and 20 of fear). These 80 recordings comprised 20 different speakers (10 male and 10 female). Participants from three separate cultures (Germany, Romania, Indonesia) were presented with forced-choice rating tasks of emotional content and recording condition ("authentic" vs. "play-acted"). Their results showed that, across all three cultures, anger was recognized more accurately when play-acted, and sadness when authentic.

Drolet and collaborators (2012) examined the authenticity of emotion expressions using an fMRI paradigm. Participants had to judge either the authenticity (authentic or play-acted) or emotional content (anger, fear, joy, or sadness) of recordings of spontaneous emotions and re-enactments by professional actors from emotional prosody stimuli. The goal of the study was to determine what specific neuronal substrates are active during judgments of emotion and authenticity, and also whether stimulus authenticity modulates the activation during these tasks. Results showed that authenticity of emotional prosody influences emotion recognition and modulates cerebral activation, and demonstrated that there was more recognition of authentic than play-acted recordings. For example, the cerebral activation for authentic stimulus was increased when compared to voluntary stimuli. Also, there was an activation of the hippocampi which can indicate an augment in the active comparison of stimuli with memories when determining authenticity (Drolet et al., 2012).

A different technique used in studies to identify authenticity includes recording spontaneous expressions of emotion through field observation (spontaneous portrayal of emotion: Anikin and Lima (2018) compared 362 acted vocalizations from seven corpora to 427 authentic vocalizations by performing an acoustic analysis to test whether vocalizations emitted spontaneously in a variety of emotionally charged situations can be distinguished from acted vocalizations intended to represent the same emotion. The authentic vocalizations were extracted from YouTube videos of people engaged in emotionally charged activities (Anikin & Lima, 2018). Results showed that the accuracy of authenticity detection varied considerably depending on the emotional category and that listeners found it relatively harder to discriminate between authentic and acted sounds of amusement (laughing), sadness (crying), and disgust when compared to authentic sounds of fear, anger and pleasure

Other authors (Dan-Glauser, 2011, McGettigan et al., 2015) choose to induce genuine emotional states in participants through the presentation of external triggers (e.g., presenting a funny video to induce laughter) or by requesting them to recall a personal memory (spontaneous portrayal of emotion) (Lima et al., 2021; Simon et al, 2008; Pinheiro et al., 2021).

Another approach used in studies on authenticity of emotions is the event-related potentials (ERPs). ERPs have a high temporal resolution in the order of milliseconds (ms), as opposed to fMRI, which presents a temporal resolution in the order of seconds (Costa, 2018). In a recent study, Conde et al (2022) used ERPs to test the possible influence of attention on the neural encoding of authenticity and also to understand how authenticity modulates the temporal course of vocal information processing in the brain. The study focused on the time course of brain responses to spontaneous and voluntary nonverbal vocalizations and also on the possibility of these responses being affected by attention. The stimuli consisted of 80 nonverbal vocalizations of laughter (40) or crying (40). Each emotional category comprised 20 voluntary and 20 spontaneous stimuli. Results suggest that authenticity and emotion cues are simultaneously and interactively processed at early stages of voice processing with authenticity properties being processed earlier for laughter than for crying.

While some studies have focused on authenticity recognition, regarding laughter (authentic vs. voluntary; McGettigan et al., 2015), others have expanded the recognition of authenticity to crying (e.g., Pinheiro et al., 2021) by presenting expressions of laughter and crying to be classified by the listeners according to authenticity, valence, arousal, trustworthiness, and dominance. Results showed that listeners could detect authenticity in both laughter and crying equally well, therefore the ability to detect authenticity in vocalizations is not limited to laughter, and that emotional authenticity shapes affective and social trait inferences from voices, with spontaneous expressions of laughter and crying being perceived as more trustworthy than the voluntary ones.

Though research in this field has been increasing, one of the main reasons for the lack of studies focusing on the recognition of authenticity in nonverbal vocalizations is the fact that the dynamic nature of vocal stimuli creates additional methodological challenges to researchers such as the combination of acoustic cues which keep changing over time (Juslin & Laukka, 2003). Hence, the use of nonverbal vocalizations as experimental stimuli (Sauter & Eimer, 2010) and techniques such as the gating paradigm could be useful to enhance our understanding of this phenomenon.

1.7 The gating paradigm

The gating paradigm provides the estimated temporal course of operations to identify the amount of acoustic-phonetic information needed to recognize an auditory stimulus. This paradigm consists of presenting stimuli in segments of increasing duration (gates), usually starting at the beginning of the stimulus and with very short duration (e.g., 20-30 ms) (Grosjean, 1980). The last gate generally corresponds to the entire stimulus. By using the gating paradigm one can estimate which is the necessary amount of acoustic information to recognize a given stimulus.

Time is crucial to be able to estimate when vocal emotions are processed by the cognitive system and end up being recognized (Schaerlaeken and Grandjean, 2018). To identify the time course of emotion recognition there is the need to identify the point in time when those emotions are identified with greater accuracy. That is called the Emotion Identification Point (EIP). The EIP, as described by Grosjean (1985), consists of identifying the gate, or segment, where the target is accurately recognized by a participant without further changes at longer gate duration for the same stimulus. In other words, the EIP indicates when the recognition of a stimuli becomes correct and stable.

The auditory gating paradigm started being used in the auditory word recognition literature (Ellis et al.,1971; Ohman, 1966; Pickett & Pollack, 1963). Grosjean (1985), for example, presented twenty monosyllabic and twenty polysyllabic nouns in context (preceded by the sentence ‘I saw’) and followed the responses to them through the next three words of the sentence (a prepositional phrase specific to each word). He could conclude that not all words are recognized before their acoustic offset, namely monosyllabic words: most of those words used in Grosjean’s study were recognized only after their offset. Research has been developing over the years as in the study by Cornew et al. (2010) using the gating paradigm with semantically devoid Jabberwocky sentences, spoken in happy, angry, and neutral intonation. The aim was to examine the recognition of emotional prosody, and its time course, and also to identify whether valence and/or arousal influence the speed and accuracy of recognizing emotion in the tone of voice. Results from the two experiments undertaken during the study showed that participants identified neutral prosody more accurately (92%) for neutral stimuli and with less auditory information compared to happy (81%) and angry (87%) prosody. Angry prosody was identified faster than happy prosody (p = .04) although contrast was not significant in the item analysis (p = .71). The mean isolation point was 403.93 ms for neutral prosody and 763.41 ms for emotional prosody. Moreover, according to the authors, accuracy, speed of processing, and error patterns favoured neutral prosody over both happy and angry prosody.

Pell and Kotz (2011) used the same paradigm, adapting it to estimate how much vocal information listeners need to be able to categorize five basic emotions (anger, disgust, fear, sadness, happiness) and neutral utterances produced by male and female English speakers. The authors stated that recognition of vocal emotion attributes in speech increases over the course of the stimulus for all tested emotions and that certain emotions can be recognized significantly better than others from the voice when evaluated in forced-choice experiments. That was the case of anger, sadness, fear, and neutral expressions between gates 1 and 4. Happiness proved different from these four emotions up to Gate 5, but with similar results on gates 6 and 7. Disgust was similar to happiness at gates 1 and 2, though being always poorly identified at the other gates. In this study, emotion identification points showed that recognition of different stimuli started, on average, in the time window of 500–600 ms (for fear, sadness, and neutral stimuli), and increased, at least, to 1000 ms (for happiness and disgust).

Another study carried out by Pell et al. (2012) using the gating paradigm, investigated the recognition of emotions from speech prosody in participants' native language and in a second language. Results showed that listeners recognized emotions reliably in both languages, though accuracy and speed were better in the native language condition, even though not for all emotions (Pell et al. 2012). Neutral, sad and especially fearful stimuli were recognized similarly in the foreign and native language conditions, whereas the recognition of angry and happy stimuli differed. Anger was identified significantly faster in English than in the foreign language condition (English: M = 356 ms; Hindi: M = 471 ms) and with happiness the opposite occurred: it was identified significantly faster in Hindi than English (English: M = 592 ms; Hindi: M = 483). The English language was more favourable for the recognition of angry stimuli while happy stimuli showed increased recognition rates in both English and Hindi at gate 6. In general, all emotions were identified between 300 and 500 milliseconds (between gates 1 and 3), except for happiness in English (around 600 milliseconds).

Jiang et al. (2015) used the same methodology to examine the nature of the in-group advantage in vocal emotion recognition. The in-group advantage, as described by Elfenbein and Ambady (2002), refers to better decoding of emotions when presented by a member of the decoder's native culture than when done by someone from a foreign culture (Elfenbein & Ambady, 2002). Jiang and colleagues compared two distinct cultures (English and Hindi) using pseudo utterances conveying four basic emotions, expressed in both languages. The pseudo utterances were divided into six gates as a function of time and not syllables. According to the results, when processing emotions in their second language, listeners are in disadvantage regarding accuracy and speed when compared to a completely foreign language. Adding to this, the in-group advantage for vocal expressions of emotion increases during speech processing after 400 ms to 500 ms of exposure to acoustic representations in speech (Jiang, 2015).

In a study by Schaerlaeken and Grandjean (2018), the gating paradigm was used to analyse the amount of information that would be necessary to decode a vocally expressed emotion. Results showed that some emotions (fear, anger, disgust) are much better recognized at full-duration than others (joy, sadness, neutral; Schaerlaeken & Grandjean, 2018). More recently, the gating paradigm was used by Castiajo & Pinheiro (2019) to study the recognition of six different nonverbal emotional vocalizations from the minimum stimulus duration. Results showed that amusement was decoded at shorter acoustic signal compared to fear, anger and sadness, opposite to other similar studies (Schröder 2003; Sauter et al. 2010). Nordstrom and Laukka (2019) undertook a study using brief utterances conveying ten emotions segmented into temporally fine-grained gates. The authors described increased emotion recognition accuracy along with increasing gate duration and stabilization after a certain duration. That is to say that anger, happiness, neutrality, and sadness were detected with above-average accuracy at the shortest or second-shortest gates (100 ms), and most other emotions were detected at gate duration of 250 ms or less, for both speech and music. EIPs were shortest for anger, joy, and sadness for both speech and music.

Regarding the gating task format, Grosjean (1980) suggested the duration-block for the experimental paradigm, implying starting the experiments always with the shortest gates and ending with the presentation of a block of all the full-duration stimuli. In this study, this procedure has been followed along with the estimation of the emotion identification point (EIP) of specific target meanings.

1.8 Objectives of the study

Some authors (Cornew et al. 2010; Pell and Kotz 2011; Rigoulot et al. 2013) have done research on duration effects on vocal emotional recognition using speech prosody stimuli. So far, to our knowledge, no studies have used the gating paradigm do examine authenticity and vocalization type recognition from nonverbal vocalizations. How long does it take for an individual to realize whether a nonverbal vocalization is authentic or voluntary? How long does it take to distinguish between laughter and crying vocalizations? In the current study, we use the gating paradigm to determine the time course of authenticity and vocalization type recognition from nonverbal vocalizations, namely those of laughter and crying. Our main goal is to determine the minimum amount of time that participants require to accurately recognize authenticity and vocalization type in the stimuli they listened to. We will then compare this time course for the recognition of authenticity with that of the type of nonverbal vocalization. This study is the first to address the time course of authenticity and its relationship with vocalization type perception from nonverbal vocalizations. The large sample size (*N* = 395) contributed to a higher reliability of and confidence on the results. Based on previous gating studies, we expect that the accuracy recognition should augment with the increase of the gate duration (e.g. Castiajo, 2019; Pell & Kotz, 2011). Regarding possible existing differences in the time course of authenticity recognition and vocalization type recognition in laughter and crying, our approach was mainly exploratory once there were no previous studies on this topic.

2. Method

2.1 Participants

The final sample comprised 395 participants (107 women, 27.1%), aged between 18 and 71 years (*M* = 26.6; *SD* = 12.0). Most participants (66.3%) had completed 12 years of education and were college students. The remaining participants had a bachelor’s degree (21.5%, *n* = 85), a master’s degree (6.3%, *n* = 25), or a PhD (4.6%, *n* = 18). One percent of the participants (*n* = 4) completed a post-graduation, and 0.3% (*n* = 1) had attended less than 9 years of education. None of the participants reported a mental or neurological disease diagnosis, or taking any medicine that would affect cognition. All were native Portuguese speakers and had normal hearing. Participants were also asked questions regarding their music training. When asked whether they had regularly practiced a musical instrument (including the voice), the majority (61.5%, *n* = 243) reported not having had any years of training; 6.3% (*n* = 25) had 1 year; 7.6% (*n* = 30) had 2 years; 4.8% (*n* = 19) had 3 years; 7.1% (*n* = 21) had 4-5 years; 7.6% (n = 30) had 6-9 years and 5.1% (*n* = 20) had 10 or more years. Similarly, when asked to complete the sentence "I have had____years of formal instruction in the practice of a musical instrument (including the voice) in my lifetime", 58.7% (*n* = 232) referred having had none; 4.6% (*n* = 18) had 6 months; 8.6% (*n* = 34) had 1 year; 9.1% (*n* = 36) had 2 years; 10.4% (*n* = 41) had 3 to 5 years; 5.6% (*n* = 22) had 6 to 9 years; and 3% (*n* = 12) had 10 or more years.

Participants were recruited through the researcher's existing professional networks, especially among university students. Informed consent was digitally obtained from all participants who read and agreed to it before starting the experiment. Participation was voluntary, and participants were not paid to take part. The study protocol was submitted to the Iscte's Ethics Committee and received approval on December 18, 2020 (reference 111/2020).

2.2 Materials - Vocal stimuli

This study used nonverbal vocalizations generated in a sound-proof anechoic chamber at University College London. Stimuli were generated by six speakers, three women and three men (aged between 24 and 48 years). These stimuli have been already used in several behavioral and neuroimaging studies (Lavan et al., 2015, 2016; Lima, Brancatisano, et al., 2016; Lima et al., 2021; O'Nions et al., 2017; Pinheiro et al., 2020). The speakers had no formal acting training but had already participated in similar tasks of vocal emotional stimuli recording, and they all stated they could perform voluntary and authentic laughter and crying when necessary.

For the authentic laughter and crying recording, emotion induction procedures were used. Speakers were asked to watch previously identified as funny video clips to record authentic laughter. Because the experimenters knew the speakers well, it was easy to promote a light environment where laughter and amusement rose naturally. As for the authentic crying recording, speakers were asked to think of actual upsetting events and were told to start by producing voluntary crying to help elicit genuine sadness.

As for the recording of voluntary laughter and crying, the six speakers were asked to simulate them without being exposed to any external eliciting stimuli, though making it sound like their vocalizations were authentic, a standard procedure generally used for the recording of acted stimuli (Amorim et al., 2021; Belin et al., 2008; Lima et al., 2013). At the end of the recordings, the six speakers stated that they felt authentic feelings of amusement and sadness during and after finishing recording authentic vocalizations, and they also referred to feelings of decreased control over their authentic vocalizations compared to voluntary ones (Pinheiro et al., 2021).

For the current study, a total of 40 stimuli were selected from a larger set of available vocalizations based on a pilot perceptual validation study (*N* = 26 listeners, none of which took part in the main study). The main criteria were to choose the same number of stimuli for each condition and that the number of stimuli should be split into male/female equally. We selected 10 stimuli from each category (voluntary crying, authentic crying, voluntary laughter, authentic laughter), five produced by men and five by women. Table 1 shows the average duration of both authentic and voluntary stimuli as used in the study for the recognition of authenticity and nonverbal vocalizations of crying and laughter.

Table 1.
Average duration in milliseconds of authentic and voluntary stimuli.

Stimuli		Average duration(ms)
Laughter	Voluntary	2.134
	Authentic	2.424
Crying	Voluntary	2.445
	Authentic	2.620

2.3 Procedure

An auditory gating task was conducted involving two judgments in each trial: one to assess emotional authenticity recognition, and the other to assess type of vocalization recognition. Listeners were provided with cumulative amounts of vocal information and asked to make judgments based on such partial vocal information.

For the construction of the gates, the 40 stimuli were first edited in version 3.0.0 of Audacity (R) software to make sure that any remaining periods of silence at the beginning and end of each stimulus were removed before the gates were created. The volume of the stimuli was also normalized for peak intensity for 70db using Praat (Boersma & Weenik, 2005), to standardize the listening experience of the different stimuli. The duration of the gates was defined as a function of time and not syllable boundaries (Pell & Kotz, 2011) because nonverbal vocalizations do not contain linguistic information. Each stimulus was segmented into eight gates and also presented in full (henceforth G100, G200, G300, G400, G650, G900, G1150, G1400, and Gfull), resulting in a total of 360 stimuli (40 stimuli x 9 durations). The first four gates were based on 100 ms increments (100, 200, 300, and 400 ms), to obtain a finer temporal resolution in the early stages of stimulus processing (Nordstrom & Laukka, 2019; see also Castiajo & Pinheiro, 2019). From the fifth gate until the eighth, stimuli were segmented every 250 ms (650, 900, 1150, and 1400 ms). Such approach to gate duration, with a more fine-grained resolution earlier in the stimuli and more coarse resolution later, has also been used in other studies (Cornew et al., 2010; Jiang et al., 2015; Nordstrom & Laukka, 2019). Stimulus presentation always started with G100 (100ms) and ended with Gfull, as usual in the literature (Grosjean, 1985; Jiang et al., 2015; Pell & Kotz, 2011), even though the order of the stimuli was randomized for each participant within each gate (i.e., laughs and cries, and authentic and voluntary stimuli, were presented randomly).

Experimental task. Because the study was conducted during the Covid-19 pandemic, the experiment was adapted for online testing using Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2018). This platform for online behavioral research allows the experiments to be performed off campus via a computer with internet connection, and all participants had that information available together with other instructions they should follow during the experiment. Data were collected between May 05 and June 23 2021. Task instructions were given in Portuguese. As participants started the experiment, they had a screen with the terms of the informed consent they should agree with to proceed. After giving consent, participants would see a welcome-to-the-task screen and, as they pressed the button 'next' at the bottom, they were directed into the trial phase of the experiment.

Instructions included explaining what participants would hear during the task: vocal expressions of emotion – laughter and crying. It was explained that those expressions were produced by several people, men and women of different ages, in different positive and negative contexts. Then, it was referred that, after listening to each sound, participants should assess whether the stimulus was a 'laugh' or a 'cry,' and whether the emotion heard was 'authentic' or 'voluntary' and what these two concepts meant: 'Authentic' indicated that the person was genuinely feeling the emotion and seemed to produce it spontaneously. 'voluntary' indicated that the person was faking the emotion and appeared to be producing it voluntarily.

Instructions also referred to the fact that participants could sometimes experience doubts about responding, for example, because the stimulus was very short (they were told that sometimes only a short part of the sound would be presented). Participants were advised to answer intuitively, in case of doubt, choosing the option that seemed closest to what they had heard.

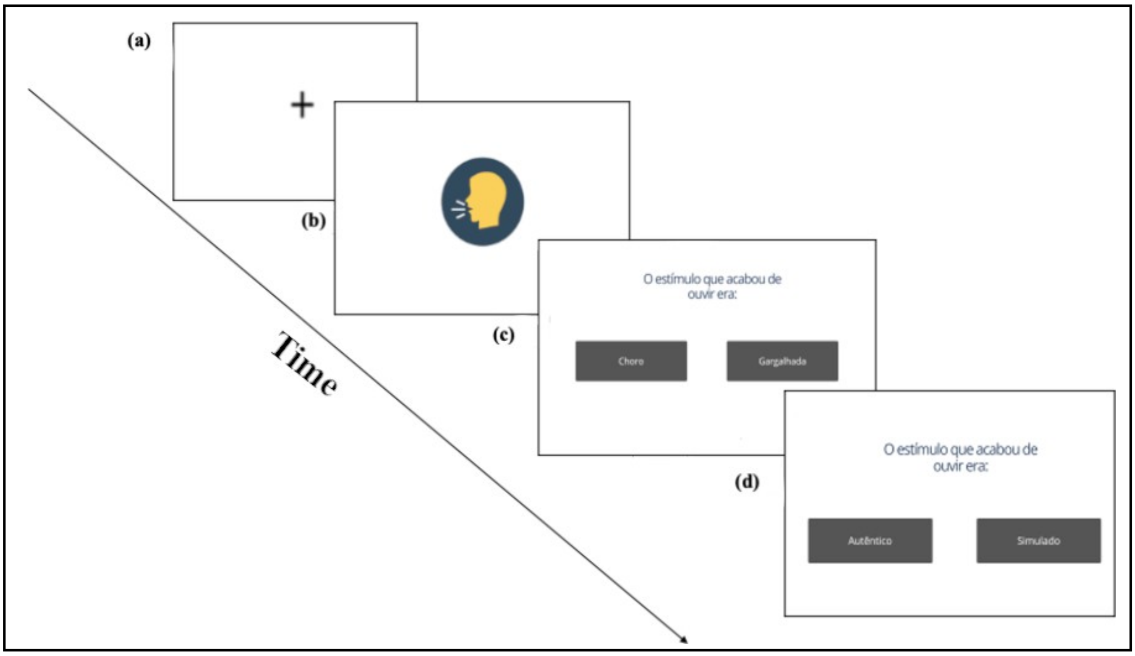
Before undertaking the task, participants were informed that they should do the experiment wearing headphones and they should be in a quiet place. In the familiarization phase, participants read that they would complete a few practice trials, to get familiarized with the task, and that they should adjust their headphones' volume to a comfortable level while doing these trials. It was also highlighted that each stimulus would only be presented once. Four practice trials were completed. The trial structure was as follows, both for the practice and experimental trials: before the stimulus was presented, a 500 ms fixation cross appeared in the middle of the screen, followed by a white screen for 100 ms, and then the stimulus was played. Participants then made two forced-choice judgement (vocalization type and authenticity), which appeared after each other in two separate screens. The order of the two judgments was randomized across participants, which means that either they had to perform a categorization of the stimulus first ('Crying' or 'Laughter') and then judge its authenticity (‘Authentic’ or ‘voluntary’), or vice-versa. There was no time limit set for the answers but the following stimulus would be displayed as soon as the participant completed the previous forced-choice judgement. Within each category, the order of the buttons in the screen was also randomized across participants (e.g., Crying – Laughter, Laughter – Crying; and the same for Authentic - voluntary).

After finishing the familiarization phase, the experiment started and went on for an average of 25 minutes. The task included listening to the 360 stimuli (divided into nine blocks, corresponding to the number of gates – 40 trials per block). Short breaks were allowed between blocks.

At the end of the task, participants were given feedback on their performance (specifically, the percentage of correct categorizations in Gfull). The structure of the trials is illustrated in Figure 1.

Figure 1.

Structure of the Gating Task



Notes. (a) Each trial starts with a fixation cross (500ms). (b) The auditory stimulus is automatically displayed right after. (c) The participant must choose if the stimulus was laughter or crying. (d) The participant must choose if the stimulus was voluntary or authentic.

Ninety-two of the participants received the Gorilla link along with all the information about the study via email or social media (e.g., WhatsApp, Facebook, Instagram) and undertook the experiment on their own. It was emphasized that they should do it in a quiet place, using headphones, and they were also asked to turn off electronic notifications while doing it. The remaining participants, corresponding to most of them (*n* = 303), were monitored via Zoom by the experimenter while performing the experiment. The experimenter could therefore ensure that all participants were indeed wearing headphones, in a quiet environment, and doing the task without interruptions.

2.4 Data analysis

We started by calculating recognition accuracy for each gate. We calculated the proportion of participants who selected each of the four vocalization categories (authentic laughter and authentic crying; voluntary laughter and voluntary crying) for each gate-duration for each recording in each task (authenticity recognition and vocalization type recognition). False alarms were also calculated (Nordstrom & Laukka, 2019).

In line with previous studies (Castiajo & Pinheiro, 2019; Pell, 2011), we analyzed the responses given by each participant to the nine gated versions of the same stimulus, from shortest to most extended gate duration (Grosjean, 1996; Pell & Kotz, 2011). In addition to these recognition accuracy data (for authenticity and type of vocalization), we also isolated the exact gate interval at which the intended target meaning was correctly identified by the participant and did not change at later gate durations for that stimulus (Castiajo & Pinheiro; 2019; Grosjean, 1996; Pell & Kotz, 2011), except if the participant identified the target emotion twice consecutively and then made only one error after that (Pell & Kotz; 2011; Salasoo & Pisoni, 1985). This dependent measure corresponds to the *identification point* (IP) of the stimulus (specified as gate 1 to 8). The IP were established separately for each stimulus when judged by each of the 395 participants.

The gate value indicating each IP in our data was individually replaced with the precise duration in milliseconds of the corresponding stimulus, such that this variable reflects the exact amount of time/acoustic information needed for authenticity and vocalization type recognition. The values in milliseconds give us the time participants needed to process vocal attributes of an utterance (Pell & Kotz, 2011). The IP for each recording was defined as the median across all participants, and the IP for each vocalization type was defined as the mean across the medians for all recordings within the same category (Nordstrom & Laukka, 2019).

We analysed the data with mixed-effects models, based on unaggregated data (i.e., data from individual trials), using R version 4.1.0, R Studio version 1.0.136, and the package lme4 version 1.1-17. We conducted models focused on two dependent measures, as per the previous gating literature: first we focused on accuracy, examining how accurately authenticity and vocalization type could be recognized at each gate duration (hit rates) (Castiajo & Pinheiro, 2019). A mixed effects model with logistic regression was estimated, considering the hit rate against the full expression as a dependent variable. The fixed factors were authenticity and vocalization type of the stimuli, and the duration of the gates (100, 200, 300, 400, 650, 900, 1150, 1400 ms and the full duration of each stimulus). A second model was used considering the same fixed factors but regarding all nine gates. Then we used a model to analyze the IP and a fourth model was used for the IP according the two tasks of authenticity recognition and vocalization type recognition.

2.5 Mixed-effects models

Accuracy for authenticity and emotion recognition were analyzed via logistic mixed-effects regression models, with random effects of participants and items. In a first set of analyses we looked for main effects of gate, authenticity, and vocalization type on authenticity and emotion discrimination accuracy. We tested separate models for each predictor, with trial-by-trial accuracy included as outcome (model tested: hit rate ~ predictor + (1|participant) + (1|stimuli).

We then tested interactions between gate and authenticity (interaction 1), gate and vocalization type (interaction 2), and authenticity and vocalization type (interaction 3) on accuracy for each task. The three models specified included trial-by-trial hit-rate as outcome and interactions 1, 2, and 3 as predictors (model tested: hit rate (each task) ~ predictor 1*predictor 2 + (1|participant) + (1|stimuli).

To analyze IPs, we conducted linear mixed-effects models with random effects for items. To explore interactions between authenticity and vocalization type for each task, models considered IP as outcome, and authenticity, vocalization type, and the interaction term as predictors (model tested: IP (each task) ~ authenticity*vocalization type + (1|stimuli).

A final model considered how triple interaction between type of task (authenticity vs vocalization type), vocalization type and authenticity might predict overall accuracy and IPs.

3. Results

3.1 Authenticity recognition

3.1.1 Recognition accuracy: overview

Figure 1 below displays the accuracy, in proportion of correct answers, for each gate, regarding the authenticity recognition task. Results for recognition accuracy are also presented in Table 1.

Figure 2.

Accuracy per gate regarding the authenticity recognition task. Average accuracy values for authentic crying and authentic laughter are shown in the solid lines. Average accuracy values for voluntary crying and for voluntary laughter are shown in the dashed lines. The grey dotted horizontal line represents the chance level (0.5).

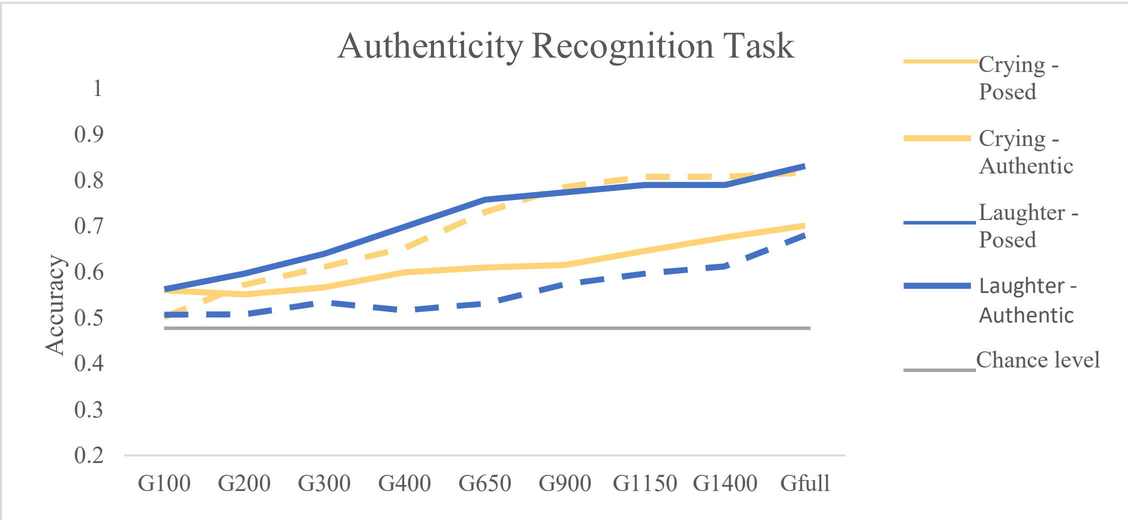


Table 2.

Mean accuracy (% target recognition) in the authenticity recognition task for all participants for each gate. Standard deviation values are given in parentheses.

Condition	Gate duration								
	G100	G200	G300	G400	G650	G900	G1150	G1400	G2000
Crying -Authentic	56.0	55.1	56.5	60.0	61.0	61.5	64.6	67.4	70.1
	(19.6)	(20.9)	(21.5)	(21.2)	(21.4)	(21.7)	(21.4)	(21.8)	(21.8)
Crying- Voluntary	50.2	57.1	61.0	65.1	73.0	78.5	80.7	80.8	81.7
	(20.7)	(20.5)	(21.9)	(22.6)	(22.1)	(20.5)	(21.2)	(20.5)	(19.4)
Laughter- Authentic	56.2	59.6	63.9	69.7	75.7	77.4	79.0	79.0	83.1
	(19.3)	(19.6)	(19.9)	(20.5)	(19.1)	(19.1)	(19.7)	(18.7)	(18.3)
Laughter- Voluntary	50.7	50.8	53.4	51.6	53.1	57.3	59.7	61.3	68.0

(21.1)	(22.4)	(21.4)	(23.2)	(23.4)	(24.7)	(25.2)	(25)	(24.8)

3.1.2 Recognition accuracy: main effects

Authenticity recognition improved with gate length, as revealed by a main effect of gate, $\chi 2$ (8) = 3725.284, $p < .001$. There were no main effects for authenticity ($p = .319$) and vocalization type ($p = .764$).

3.1.3 Recognition accuracy: interaction effects

A significant interaction between gate and authenticity, $\chi 2$ (8) = 45.086, $p < .001$, reveals that recognition of authentic and voluntary stimuli differed across gates. Gate and type of vocalization also interacted, $\chi 2$ (8) = 30.121, $p < .001$, showing that authenticity recognition differed across gates depending on the type of vocalization (laughter vs crying).

Authenticity recognition was also modulated by authenticity and type of vocalization, $\chi 2$ (8) = 45.086, $p < .001$, confirming that authenticity identification differed across the authenticity and vocalization type categories. Post-hoc comparisons show that, for authentic vocalizations, accuracy was higher for laughter than for crying (estimated difference = 0.111, *SE* = 0.041, z = 2.700, $p = .007$), while the inverse pattern was observed for voluntary stimuli: accuracy was higher for crying than for laughter (estimated difference = 0.139, *SE* = 0.043, $z = 3.239$, $p = .002$).

3.1.4 Recognition accuracy and chance level

We used t-tests to compare the average recognition of authenticity along the gates with the chance level and to understand which was the shortest gate at which authenticity was significantly identified above the chance level for each vocalization type. Results showed that authentic crying $t(395) = 7.628$, $p < .001$ and authentic laughter $t(395) = 7.889$, $p < .001$ were recognized above chance level from the shortest gate (G100) onwards. Results also showed that voluntary crying was only recognized above chance level from G200, $t(395) = 9.066$, $p < .001$, and voluntary laughter from G650, $t(395) = 3.889$, $p < .001$.

3.1.5 Authenticity Identification Points

After analyzing accuracy for authenticity, we determined the identification points (IPs) for each task. The IPs (specified as gate G100 to G1400) were determined for each stimulus and calculated separately for authenticity recognition and vocalization type recognition for each of the 395 participants. Table 2 presents the distribution of IPs for each vocalization type at each of the eight gates for the authenticity recognition task.

3.1.6 Authenticity IPs: Interaction effects

As presented on Table 2, authentic laughter and voluntary crying were the conditions with the most IPs at shorter gates (G100) regarding the authenticity recognition task. Statistical analyses did not reveal a main effect of vocalization type, F (1, 35.64, N = 395) = .112, p = .739, or authenticity, F (1, 35.64) = 0.367, p = .549. However, a significant interaction between authenticity and vocalization type was observed, F (1, 35.64) = 13.341, $p < .001$. Follow up comparisons showed that participants relied on less acoustic information to accurately discriminate authentic laughter than voluntary crying (estimated difference = -162.966, SE = 57.783, z = -2.82, p = .010). The reverse pattern was also found, i.e., participants relied on less acoustic information to recognize authentic crying compared to voluntary laughter (estimated difference = -135.561, SE = 57.802, z = -2.34, p = .019). These results follow the same pattern previously found for accuracy ratings (better recognition of authentic laughter and voluntary crying).

Table 3.

Distribution of correct and incorrect identification points (IPs) for each condition and gate for the authenticity recognition task. Total incorrect corresponds to the stimuli within each condition in which participants did not stabilize their response.

Condition	Gate duration								Total Correct	Total Incorrect
Gate	G100	G200	G300	G400	G650	G900	G1150	G1400		
Crying - Authentic	546	223	324	302	291	337	341	294	2658	1292
	(13.8%)	(5.7%)	(8.2%)	(7.7%)	(7.4%)	(8.5%)	(8.6%)	(7.4%)	(67.3%)	(32.7%)
Crying - Voluntary	747	459	455	401	417	362	277	162	3280	670
	(18.9%)	(11.6%)	(11.5%)	(10.1%)	(10.6%)	(9.2%)	(7%)	(4.1%)	(83%)	(17%)
Laughter - Authentic	934	472	459	376	341	239	199	172	3192	758
	(23.7%)	(11.9%)	(11.6%)	(9.5%)	(8.6%)	(6%)	(5%)	(4.4%)	(80.8%)	(19.2%)
Laughter - Voluntary	496	303	272	204	269	298	303	288	2433	1517
	(12.6%)	(7.7%)	(6.9%)	(5.2%)	(6.8%)	(7.5%)	(6.7%)	(7.3%)	(61.6%)	(38.4%)

Note. IPs are based on a maximum of 15800 observations (395 participants * 40 items).

Table 2 shows that the concentration of correct answers for authentic laughter at the shortest gate is higher than for the other three conditions, which means that stimuli from this condition are more accurately recognized from the beginning and stabilize much earlier than the remaining stimuli. It also shows that G100 is the gate where more stimuli stabilize, and this is a true for all conditions.

3.2 Vocalization type recognition

3.2.1 Recognition accuracy: overview

Figure 3 displays the accuracy, in proportion of correct answers, for each of the gates, considering the vocalization type and authenticity represented by the stimuli. The percentage of recognition accuracy is shown in Table 3.

Figure 3.

Accuracy per gate regarding the vocalization type recognition task. Average accuracy values for authentic crying and authentic laughter are shown in the solid lines. Average accuracy values for voluntary crying and for voluntary laughter are shown in the dashed lines. The grey dotted horizontal line represents the level of chance (0.5).

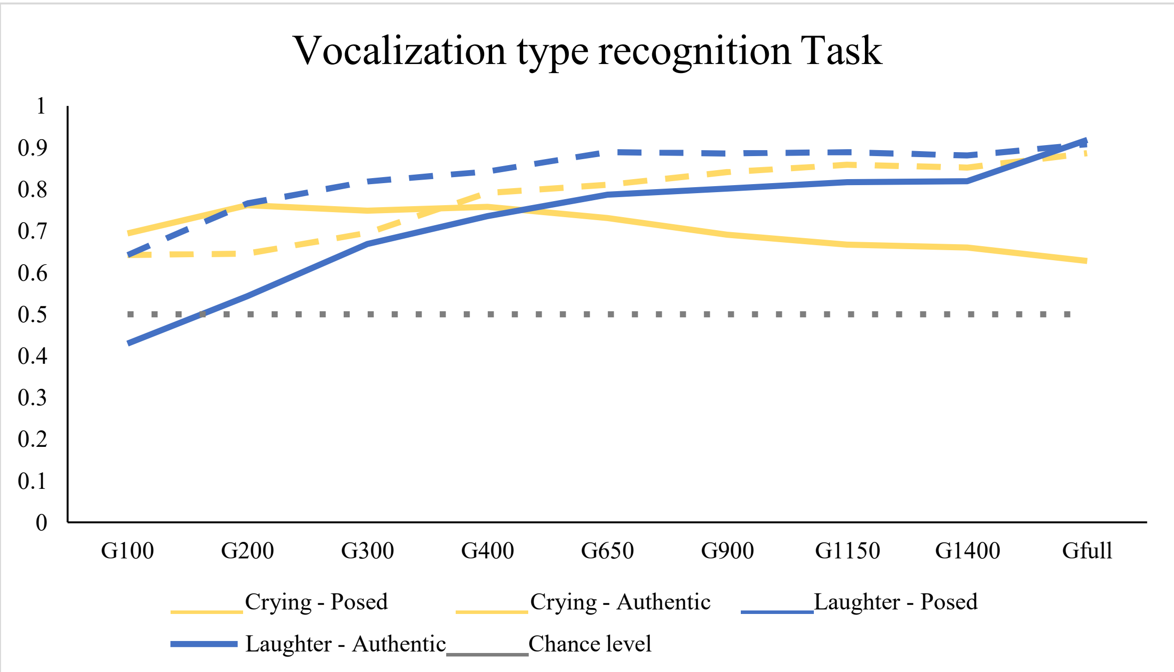


Table 4.

Average of accuracy (% target recognition) for the vocalization type recognition task for all participants for each gate. Standard deviation values are given in parentheses.

Condition	Gate duration								
	G100	G200	G300	G400	G650	G900	G1150	G1400	Gfull
Crying-Authentic	69.5	76.2	74.9	75.8	73.1	69.2	66.6	65.9	62.8
	(19)	(19.6)	(18.8)	(19.5)	(18.9)	(18.1)	(18.2)	(17.2)	(17)
Crying- Voluntary	64.2	64.6	69.6	79.2	81.2	84.2	85.9	85.3	88.7
	(18.7)	(17)	(17.9)	(19.6)	(19.2)	(18.2)	(18.7)	(19.3)	(17.3)
Laughter-Authentic	42.9	54.3	66.8	73.5	78.7	80.2	81.7	81.9	91.9
	(18.4)	(16.6)	(15.3)	(14.6)	(14.4)	(15)	(15.4)	(14.4)	(14.4)
Laughter- Voluntary	64.2	76.6	82.0	84.4	89.0	88.7	88.9	88.2	90.9
	(21.1)	(16.8)	(15.7)	(17.3)	(16.2)	(17)	(18.3)	(17.9)	(14.1)

3.2.2 Vocalization type recognition: main effects

Participant's accuracy improved with gate length, $\chi2$ (8) = 4647.61, *p* <.001. There was also a marginal effect of authenticity type, $\chi2$ (1) = 3938, *p* = .047, showing that participants were slightly better at categorizing voluntary stimuli compared to authentic stimuli. Vocalization type identification accuracy was similar across vocalizations, $\chi2$ (1) = 1.000, *p* = .317.

3.2.3 Vocalization type recognition: interaction effects

Results showed a significant interaction between gate and authenticity, $\chi2$ (8) = 334.641, *p* < .001, revealing that vocalization type accuracy differed across gates, with higher hit rates for the voluntary stimuli versus the authentic ones across all gates. Gate also interacted with vocalization type, $\chi2$ (8) = 1962.637, *p* < .001, indicating that accuracy was different across gates depending on the vocalization type being expressed (laughter vs crying). Interactions between authenticity and type of vocalization were non-significant, $\chi2$ (8) = 0.313, $p = .576$), indicating that vocalization type identification was similarly accurate for both authentic and voluntary stimuli.

3.2.4 Recognition accuracy and chance level

Similarly to the authenticity recognition task, we used t-tests to compare the average recognition of each type of vocalization over the gates with the chance level and to understand which was the shortest gate at which each vocalization type was significantly identified above the chance level. Results showed that all conditions were recognized above chance level at the first gate (G100; ps < .001), with the exception of authentic laughter ($t(395) = -8.967$, $p = 1.000$),

3.2.5 Vocalization type Identification Points

The process described in Section 1.3. was also followed to determine IPs for vocalization type. Table 4 presents the distribution of IPs for each category per gate interval.

Table 4 suggests that the categories with the most IP at shorter gates (G100) were voluntary laughter and voluntary crying, with the concentration of IPs being much lower in subsequent gates. Authentic crying also reached higher results during the first gates, although numbers decreased from the third gate (G300) onwards. Authentic laughter was the condition with lowest IPs throughout all gates.

Table 5.

Distribution of correct and incorrect identification points (IPs) for each condition and gate for the vocalization type recognition task. Total incorrect corresponds to the stimuli within each condition in which participants did not stabilize their response.

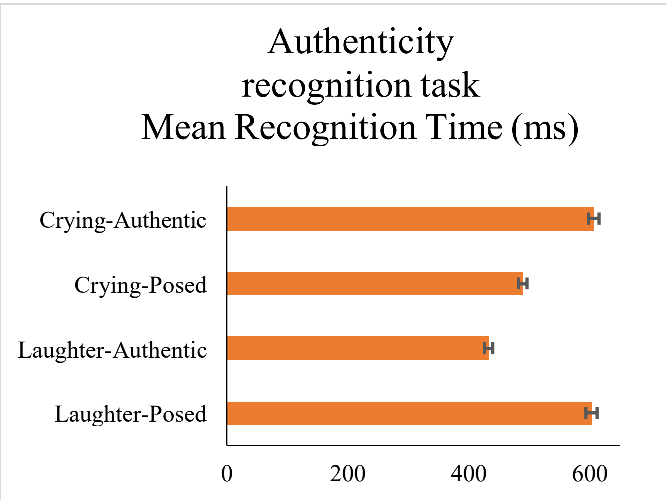
Condition	Gate duration (ms)							Total Correct	Total Incorrect
Gate	G100	G200	G300	G400	G650	G900	G1150	G1400		
Crying-Authentic	33.2	8.5	6.2	3.9	3.5	3.4	3.6	2.4	2561	1389
	(33.2%)	(8.5%)	(6.2%)	(3.9%)	(3.5%)	(3.4%)	(3.6%)	(2.4%)	(64.8%)	(35.2%)
Crying - Voluntary	1389	392	514	418	277	201	187	135	3513	437
	(35.2%)	(9.9%)	(13%)	(10.6%)	(7%)	(5%)	(4.7%)	(3.4%)	(88.9%)	(11.1%)
Laughter - Authentic	866	777	685	377	241	133	109	132	3320	630
	(21.9%)	(19.7%)	(17.3%)	(9.5%)	(6.1%)	(3.4%)	(2.8%)	(3.3%)	(84.1%)	(15.9%)
Laughter - Voluntary	1763	664	391	265	240	110	96	81	3610	340
	(44.6%)	(16.8%)	(9.9%)	(6.7%)	(6.1%)	(2.8%)	(2.4%)	(2.1%)	(91.4%)	(8.6%)

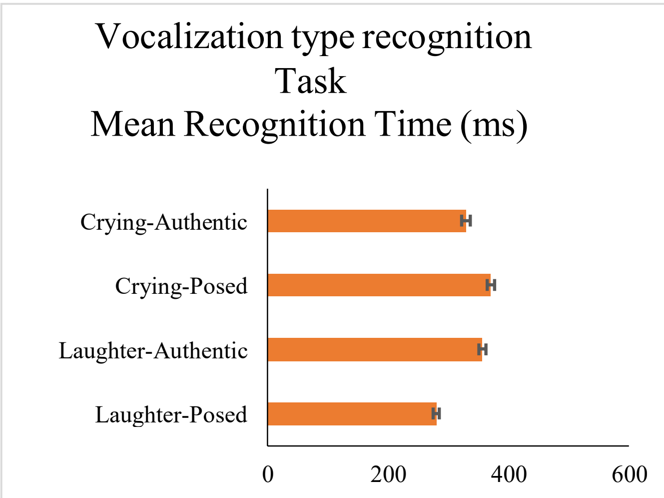
Note. IPs are based on a maximum of 15800 observations (395 participants X 40 items).

Figures 4 and 5 below show the average IPs for each vocalization type

Figure 4. and Figure 5.

Identification points (IPs) for each condition during the authenticity recognition task (Fig 4) and the vocalization type recognition task (Fig 5). Values of IP presented correspond to the average of IPs for the stimuli corresponding to the four conditions: authentic crying, voluntary crying, authentic laughter, voluntary laughter. Error bars refer to standard error.





By looking at Figure 4, it seems that for the authenticity recognition task, participants rely on less acoustic information to identify authentic crying and voluntary laughter. Results, however, do not support this assumption, as evidenced by the absence of main effects for vocalization type (p = .595) and authenticity (p = .165), as well as a non-significant interaction between vocalization type and authenticity (p = .280). On the other hand, Figure 5 shows that voluntary crying and authentic laughter were identified more often faster than the other two conditions, during the vocalization type recognition task.

4. Discussion

No previous studies have focused on the time course of authenticity perception and its relation with the time course of nonverbal emotional vocalizations. This fact could be taken as a weakness in terms of lack of solid ground for drawing conclusions. However, it can also be referred to as the major strength of the present study, as it triggers the possibility of originating future studies on this topic. Also, the fact that the gating paradigm was used as the main tool to infer about the time course of authenticity must be referred to as added-value once it allows a precise and accurate measure to determine the time course of authenticity by indicating the exact amount of acoustic information needed to identify a stimulus (Grosjean, 1996).

The primary objective of this study was to determine the precise amount of acoustic information people need to accurately recognize authenticity in nonverbal emotional vocalizations, and also to determine the time course needed to accurately identify between two different types of nonverbal emotional vocalizations (laughter or crying).

Regarding the four types of vocalizations used in this study, during the authenticity recognition task, results showed that authentic stimuli were identified faster than voluntary stimuli and above chance level from the first gate (G100) onwards. Voluntary vocalizations were identified above chance level after G200 (voluntary crying) and G650 (voluntary laughter). Adding to this, the concentration of more correct answers in the first gate (G100) regarding the identification of authentic laughter is higher and stabilizes much before the identification of the other three types of vocalizations. Also, we can refer to a privileged detection of authenticity markers in authentic laughter (Neves et al., 2018).

As for the vocalization type recognition task, results showed that voluntary vocalizations are identified faster than the authentic ones and with the exception of authentic laughter, all conditions were recognized above chance level at the first gate (G100).

The approach of the present study was mainly exploratory once previous literature on the time course of authenticity and vocalization type recognition is scarce. Some authors had already focused on the study of authenticity of emotions, but using different procedures, such as those used by Drolet through fMRI paradigms (Drolet et al., 2012), or event-related potentials (Conde et al., 2022) for a better understanding of the role of authenticity in modulating the temporal course of vocal information processing in the brain. On another note, Anikin and Lima (2018) tested whether vocalizations emitted spontaneously can be distinguished from acted vocalizations even when representing the same emotion. In line with Lavan and colleagues (2017), who aimed at distinguishing if emotions are spontaneously or voluntarily expressed in daily life, we aimed further, using the gating paradigm to understand how much time and acoustic information people need to make that distinction effectively.

4.1 Authenticity recognition

Results from previous studies (McGettigan et al., 2015; Pinheiro et al., 2021) showed that listeners could detect authenticity in both laughter and crying equally well. Our results suggest that the perception of laughter and crying depended on the authenticity of the emotional stimuli. As far as authentic emotional vocalizations are concerned, participants seem to be faster at identifying laughter than crying. When it comes to voluntary emotional vocalizations, participants were faster at identifying crying.

Our results confirmed our prediction that authenticity recognition of authentic stimuli increased with gate duration, regardless of the type of vocalization. However, authentic laughter was identified at shorter gates compared to authentic crying. Regarding voluntary stimuli, the recognition accuracy for voluntary crying happened at earlier stages than for voluntary laughter, in line with previous studies using ERPs (Conde et al, 2022) and EEG (Kosilo et al., 2021). In fact, recognition accuracy percentages for authenticity of authentic laughter and voluntary crying are very similar and both represent the conditions with the highest identification points at shorter gates.

4.2 Vocalization type recognition

The results of this study demonstrate that the time course for recognizing the vocalization type (laughter vs. crying) is shorter when the vocalizations are voluntary (due to possible standardization) than when vocalizations are authentic. Some studies suggest that authentic laughter, for example, is extremely variable regarding time and production mode (Bachorowskia et al. 2001) and more extreme than voluntary laughter. That is, being longer in duration, less voiced, higher-pitched, and with higher spectral centre of gravity and intensity (Lavan et al, 2012). These two factors, variability and being extreme, might have influenced the participants in terms of the time course. Our results demonstrated that the vocalizations with the highest identification points at shorter gates were voluntary laughter and voluntary crying. These findings complement previous studies such as the study by Castiajo and Pinheiro, which results showed that the identification of amusement and sadness happened at shorter acoustic signals than the identification of other nonverbal vocalizations.

4.3 Identification Points

We used identification points to be able to accurately identify the exact point in time, measured by a gate interval, when a stimulus is recognized and its identification does not change from then on (Pell et al., 2011). In the authenticity recognition task, results showed that less acoustic information is necessary to identify authentic laughter compared to voluntary crying, in line with previous studies referring to authentic laughter as reflecting a less controlled and genuinely felt emotion (Neves et al., 2018). Some authors mention that the early identification of laughter might be connected with its social significance and acoustic distinctiveness in nonverbal social communication (Castiajo and Pinheiro (2019). Inversely, less acoustic information is necessary to recognize authentic crying compared to voluntary laughter.

Based on previous gating studies, we predicted that the accuracy recognition should improve with the increase of the gate duration (Pell & Kotz, 2011; Castiajo, 2019; Nordstrom & Laukka, 2019). According to some authors, the identification points for sadness and happiness tend to happen at different stages, with sadness being recognized at earlier stages than happiness (Schröder 2003; Sauter et al. 2010; Pell & Kotz, 2011). On the other hand, Castiajo and Pinheiro (2019) proved otherwise with amusement being decoded at a shorter acoustic signal compared to sadness. Our results are in line with these studies and also suggest that authentic laughter was significantly better identified than the other three conditions, maybe due to a so-called social bias in which laughter is considered as more authentic and spontaneous than crying (Anikin and Lima, 2018; Conde et al, 2022; Pinheiro et al., 2021) and something that we can use (not necessarily knowingly) to establish and maintain social bonds (Scott et al., 2014)

4.4 Limitations

The first limitation might have to do with accuracy in terms of the results. Studies on the time course of emotional authenticity processing using event-related potentials (Conde et al., 2022) instead of a gating paradigm might have more accurate results, because ERPs can identify how much time it the brain needs from stimulus onset until a response is made. On the other hand, the gating paradigm deals with the evaluation processes that lead to a response. However, event-related potentials are not as accurate as the gating paradigm in terms of identifying the time course of vocal information needed for above-chance recognition of nonverbal emotional vocalizations (Castiajo & Pinheiro, 2019). Another limitation is related with the fact that the experiment was performed online. Although participants were supervised via zoom throughout the experiment and were asked about potential problems occurred during the experimental session, this did not fully guarantee that participants were focused on the task or that the experiment was successfully completed. It is also important to note that the length of the experiment might have influenced the participants performance due to fatigue or boredom. Another limitation was the impossibility of using the same sound system (comprising the computer and ear/headphones) for all the participants: some referred to not having had the possibility of listening to all sounds in a clear way. Also, the fact that most participants were men can be a weakness, in the sense that results would be more reliable if drawn from a balanced group of participants regarding sex, according to the literature on the subject of gender differences in identifying emotions (Wells et al., 2016;

Wingenbach et al. 2018). Namely, because according to some authors those differences account for female advantage in decoding accuracy in vocal emotions (Lausen and Schacht, 2018).

5. Conclusion

The current study adds to emotional prosody research by shedding light on the time required to decode authenticity in nonverbal vocalizations. Results showed that we are faster at identifying the type of vocalization (laughter vs. crying) than at identifying their authenticity (authentic vs. voluntary). Still, it is worth referring that authentic stimuli were identified faster than voluntary stimuli. Adding to this, authentic laughter identification is higher and stabilizes much before the identification of the other three types of vocalizations. Regarding vocalization type recognition, voluntary nonverbal vocalizations (both laughter and crying) were identified with less acoustic information than authentic stimuli.

6. References

Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinheiro, A. P. (2021). Changes in vocal emotion recognition across the life span. *Emotion*, *21*(2), 315. https://psycnet.apa.org/doi/10.1037/emo0000692

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Quarterly Journal of Experimental Psychology, 71*(3), 622–641. https://doi.org/10.1080/17470218.2016.1270976

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N. & Evershed, J. (2018). Gorilla in our Midst: An online behavioral experiment builder. *bioRxiv*, 438242. https://doi.org/10.1101/438242

Audacity Team (2012) Audacity®. Version 2.0.0. Audio editor and recorder. Available from: http://audacityteam.org/ (Accessed 03/2021).

Bachorowski, J. A., Smoski, M. J., & Owren, M. J. (2001). The acoustic features of human laughter. *The Journal of the Acoustical Society of America, 110*(3), 1581-1597. DOI: 10.1121/1.1391244

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. Emotion, *12*(5), 1161. https://psycnet.apa.org/doi/10.1037/a0025827

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*, 531-539.

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in cognitive sciences, 8*(3), 129-135. https://doi.org/10.1016/j.tics.2004.01.008

Blasi, A., Lloyd-Fox, S., Sethna, V., Brammer, M. J., Mercure, E., Murray, L., & Johnson, M. H. (2015). Atypical processing of voice sounds in infants at risk for autism spectrum disorder. *Cortex, 71*, 122-133. https://doi.org/10.1016/j.cortex.2015.06.015

Bylsma L. M., Gračanin A., Vingerhoets A. J. (2019). The neurobiology of human crying. *Clinical Autonomic Research*, *29*, 63–73. https://doi:10.1007/s10286-018-0526-y

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.51, Retrieved July 22, 2021, from http://www.praat.org/

Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior*, *35*(4), 327-335. https://doi.org/10.1016/j.evolhumbehav.2014.03.003

Bryant, G. A., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture, 8*(1-2), 135–148. https://doi.org/10.1163/156770908X289242

Castiajo, P., & Pinheiro, A. (2019). Decoding emotions from nonverbal vocalizations: How much voice signal is enough? *Motivation and Emotion, 43*(5), 803-813. https://doi.org/10.1007/s11031-019-09783-9

Cornew, L., Carver, L. & Love, T. (2010) There's more to emotion than meets the eye: A processing bias for neutral content in the domain of emotional prosody, *Cognition and Emotion, 24(*7), 1133-1152, DOI: 10.1080/02699930903247492

Costa, M. G. B. (2018). Authenticity recognition in laughter and crying: an ERP study (Doctoral dissertation).

Cowen, A.S., Laukka, P., Elfenbein, H.A. et al. (2019). The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Natural Human Behavior, 3*, 369–382 . https://doi.org/10.1038/s41562-019-0533-6

Dan-Glauser, E.S., Scherer, K.R (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research, 43,* 468. https://doi.org/10.3758/s13428-011-0064-1

Davila-Ross, M., Allcock, B., Thomas, C., & Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion, 11*(5), 1013. https://psycnet.apa.org/doi/10.1037/a0022594

Drolet, M., Schubotz, R. I., & Fischer, J. (2014). Recognizing the authenticity of emotional expressions: F0 contour matters when you need to know. *Frontiers in human neuroscience, 8*, 144. https://doi.org/10.3389/fnhum.2014.00144

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence. *Cognitive, affective & behavioral neuroscience, 12*(1), 140–150. https://doi.org/10.3758/s13415-011-0069-3

Ekman, P., & Keltner, D. (1997). Universal facial expressions of emotion. Segerstrale U, P. Molnar P, eds. *Nonverbal communication: Where nature meets culture, 27*, 46.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*, 203-235. http://dx.doi.Org/10.1037/0033-2909.128.2.203

Ellis, L., Derbyshire, A. J., & Joseph, M. E. (1971). Perception of electronically gated speech. *Language and speech, 14*(3), 229-240. https://doi.org/10.1177%2F002383097101400303

Gervais, M.M., & Wilson, D.S. (2005). The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly Review of Biology, 80*, 395 - 430. http://dx.doi.org/10.1086/498281

Gorilla Experiment Builder, https://app.gorilla.sc

Grosjean, F., & Frauenfelder, U. H. (1996). A guide to spoken word recognition paradigms: Introduction. *Language and Cognitive Processes, 11*(6), 553-558. https://doi.org/10.1080/016909696386935

Grosjean, F. (1996) Gating, Language and Cognitive Processes, *11*(6), 597-604, DOI: 10.1080/016909696386999

Grosjean, F. (1985)The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics, 38*, 299–310. https://doi.org/10.3758/BF03207159

Grosjean F. (1980) Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*(4):267–283. https://doi.org/10.3758/BF03204386

Jiang, X., Paulmann, S., Robin, J., & Pell, M. D. (2015). More than accuracy: Nonverbal dialects modulate the time course of vocal emotion recognition across cultures. *Journal of Experimental Psychology: Human Perception and Performance, 41*(3), 597–612. https://doi.org/10.1037/xhp0000043

Johnson J. A. (2017). Speech, Voice, and Communication. *International review of neurobiology, 134*, 1189–1205. https://doi.org/10.1016/bs.irn.2017.04.014

Juslin, P.N., Laukka, P. & Bänziger, T. (2018).The Mirror to Our Soul? Comparisons of Spontaneous and Posed Vocal Expression of Emotion. *Journal of Nonverbal Behaviour 42*, 1–40. https://doi.org/10.1007/s10919-017-0268-x

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129,* 770–814. https://doi:10.1037/0033-2909.129.5.770

Kosilo, M., Costa, M., Nuttall, H.E. et al. (2021).The neural basis of authenticity recognition in laughter and crying. *Scientific Reports 11*, 23750 https://doi.org/10.1038/s41598-021-03131-z

Laukka, P., Elfenbein, H. A., Thingujam, N. S., Rockstuhl, T., Iraki, F. K., Chui, W., & Althoff, J. (2016). The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features. *Journal of personality and social psychology, 111*(5), 686–705. https://doi.org/10.1037/pspi0000066

Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Chui, W., Iraki, F. K., Rockstuhl, T., & Thingujam, N. S. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in psychology, 4*, 353. https://doi.org/10.3389/fpsyg.2013.00353.

Lausen, A., & Schacht, A. (2018). Gender Differences in the Recognition of Vocal Emotions. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.00882

Lavan, N., & McGettigan, C. (2017). Increased discriminability of authenticity from multimodal laughter is driven by auditory information. *Quarterly journal of experimental psychology, 70*(10), 2159–2168. https://doi.org/10.1080/17470218.2016.1226370

Lavan, N., Rankin, G., Lorking, N., Scott, S., & McGettigan, C. (2017). Neural correlates of the affective properties of spontaneous and volitional laughter types. *Neuropsychologia, 95*, 30-39.

Lavan, N., Scott, S.K. & McGettigan, C. (2016). Laugh Like You Mean It: Authenticity Modulates Acoustic, Physiological and Perceptual Properties of Laughter. *Journal of Nonverbal Behavior, 40*(2), 133–149 (2016). https://doi.org/10.1007/s10919-015-0222-8

Lavan, N., Lima, C. F., Harvey, H., Scott, S. K., & McGettigan, C. (2015). I thought that I heard you laughing: Contextual facial expressions modulate the perception of authentic laughter and crying. *Cognition and Emotion, 29*(5), 935–944. https://doi.org/10.1080/02699931.2014.957656

Lima, C., Arriaga, P., Anikin, A., Pires, A. R., Frade, S., Neves, L., Scott, K. S. (2021), Authentic and posed emotional vocalizations trigger distinct facial responses, *Cortex*, *141*, 280-292. DOI: https://doi.org/10.1016/j.cortex.2021.04.015

Lima, C. F., Brancatisano, O., Fancourt, A., Mullensiefen, D., Scott, S. K., Warren, J. D., & Stewart, L. (2016). Impaired socioemotional processing in a developmental music disorder. *Scientific Reports, 6*, 34911. https://doi.org/10.1038/srep34911

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods, 45*(4), 1234-1245. https://doi.org/10.3758/s13428-013-0324-3.

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2015). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral cortex, 25*(1), 246-257. https://doi.org/10.1093/cercor/bht227

McKeown, G., Sneddon, I., & Curran, W. (2015). Gender differences in the perceptions of genuine and posed laughter and amused facial expressions. *Emotion Review, 7*(1), 30–38. https://doi.org/10.1177%2F1754073914544475

Neves, L., Cordeiro, C., Scott, S. K., Castro, S. L., & Lima, C. F. (2018). High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. *Quarterly Journal of Experimental Psychology, 71*(11), 2355–2363. https://doi.org/10.1177/1747021817741800

Nordström, H. (2019). Emotional Communication in the Human Voice (PhD dissertation, Department of Psychology, Stockholm University). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-167973.

Nordström, H., & Laukka, P. (2019). The time course of emotion recognition in speech and music. *The Journal of the Acoustical Society of America, 145*(5), 3058. https://doi.org/10.1121/1.5108601

Nummenmaa, L., & Calvo, M. G. (2015). Dissociation between recognition and detection advantage for facial expressions: a meta-analysis. *Emotion, 15*(2), 243. https://psycnet.apa.org/doi/10.1037/emo0000042

Öhman, S. E. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America, 39*(1), 151-168. https://doi.org/10.1121/1.1909864.

O'Nions, E., Lima, C. F., Scott, S. K., Roberts, R., McCrory, E. J., & Viding, E. (2017). Reduced laughter contagion in boys at risk for psychopathy. *Current Biology, 27*(19), 3049-3055. https://doi.org/10.1016/j.cub.2017.08.062.

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological psychology, 111,* 14-25. https://doi.org/10.1016/j.biopsycho.2015.08.008

Pell, M. D., Robin, J., & Paulmann, S. (2012). How quickly do listeners recognize emotional prosody in their native versus a foreign language?. *Proc. Speech Prosody*. 2012, 665-668.

Pell, M. D., Kotz SA (2011) On the Time Course of Vocal Emotion Recognition. *PLoS ONE 6*(11): e27256. https://doi.org/10.1371/journal.pone.0027256

Pinheiro, A., Anikin, A., Conde, T., Sarzedas, J., Chen, S., Scott, S., & Lima, C. (in print) (2021) Emotional authenticity modulates affective and social trait inferences from voices. *Philosophical Transactions R. Soc. B* 3762020040220200402 http://doi.org/10.1098/rstb.2020.0402

Pires, A. R. P. (2019). Contágio emocional e deteção de autenticidade em vocalizações não verbais: Estudo comportamental e psicofisiológico [Dissertação de mestrado, Iscte-Instituto Universitário de Lisboa]. Repositório Iscte. https://hdl.handle.net/10071/19413

Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016 Voice Modulation: A Window into the Origins of Human Vocal Control? *Trends Cogn. Sci. 20*, 304–318. https://doi:10.1016/j.tics.2016.01.002

Pollack, I., & Pickett, J. M. (1963). The Intelligibility of Excerpts from Conversation. *Language and Speech, 6*(3), 165–171. https://doi.org/10.1177/002383096300600305

Provine, R. R., Krosnowski, K. A., & Brocato, N. W. (2009). Tearing: Breakthrough in human emotional signaling. *Evolutionary Psychology, 7*(1), 147470490900700107. https://doi.org/10.1177%2F147470490900700107

Rigoulot, S., Wassiliwizky, E., & Pell, M. D. (2013). Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition. *Frontiers in psychology, 4,* 367. https://doi.org/10.3389/fpsyg.2013.00367.

Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of memory and language, 24*(2), 210-231. https://doi.org/10.1016/0749-596X(85)90025-7

Sauter, D. A. (2017). The nonverbal communication of positive emotions: An emotion family approach. *Emotion Review, 9*(3), 222-234. https://doi.org/10.1177%2F1754073916667236

Sauter, D. A., & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *Journal of cognitive neuroscience, 22*(3), 474–481. https://doi.org/10.1162/jocn.2009.21215

Sauter, Eisner, Ekman, et al. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. Proceedings of the *National Academy of Sciences of the United States of America, 107*(6), 2408–2412. https://doi:10.1073/pnas.0908239106

Sauter, D. (2007) An investigation into vocal expressions of emotions: the roles of valence, culture, and acoustic factors. Doctoral thesis , University of London. https://discovery.ucl.ac.uk/id/eprint/1445045

Schaerlaeken, S., & Grandjean, D. (2018). Unfolding and dynamics of affect bursts decoding in humans. *PloS One, 13*(10), e0206216. https://doi.org/10.1371/journal.pone.0206216

Scheiner, E., & Fischer, J. (2011). Emotion expression: The evolutionary heritage in the human voice. In W. Welsch (Ed.), *Interdisciplinary anthropology: The continuing evolution of man,* 105-129. Heidelberg, Germany: Springer. https://doi.org/10.1007/978-3-642-11668-1_5

Scherer, K.R., Bänziger, T. (2004) Emotional expression in prosody: a review and an agenda for future research. *Proceedings Speech Prosody,* 359-366

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76–92. https://doi.org/10.1177/0022022101032001009

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

Scott, S.K., Sauter, D.A., & McGettigan, C. (2009). Brain mechanisms for processing perceived emotional vocalizations in humans. *Handbook of Behavioral Neuroscience, 19*, 187-198. https://doi:10.1016/B978-0-12-374593-4.00019-X

Scott, S. K., Lavan, N., Chen, S., & McGettigan, C. (2014). The social life of laughter. *Trends in cognitive sciences, 18*(12), 618–620. https://doi.org/10.1016/j.tics.2014.09.002

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2013). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *Pain, 135*, 55–64. https://doi:10.1016/j.pain.2007.05.008

Simons, G., Bruder, M., van der Löwe, I., & Parkinson, B. (2013). Why try (not) to cry: intra- and inter-personal motives for crying regulation. *Frontiers in psychology, 3*, 597. https://doi.org/10.3389/fpsyg.2012.00597

Titze, I.R. (2000), Principles of Voice Production. National Center for Voice and Speech. ISBN 13: 9780874141221.

Thompson, W. & Balkwill, L. (2006). Decoding speech prosody in five languages, *2006*(158), 407-424. https://doi.org/10.1515/SEM.2006.017

Vasconcelos, M., Dias, M., Soares, A. P., & Pinheiro, A. P. (2017). What is the melody of that voice? Probing unbiased recognition accuracy with the Montreal affective voices. *Journal of Nonverbal Behavior, 41*(3), 239-267. https://psycnet.apa.org/doi/10.1007/s10919-017-0253-4

Vingerhoets, A. (2013). Why only humans weep: Unravelling the mysteries of tears. Oxford University Press. ISBN: 0198570244, 9780198570240

Vingerhoets, A., & Bylsma, L. M. (2016). The riddle of human emotional crying: A challenge for emotion researchers. *Emotion Review, 8*(3), 207-217. https://doi.org/10.1177/1754073915586226

Wells, L. J., Gillespie, S. M., and Rotshtein, P. (2016). Identification of emotional facial expressions: effects of expression, intensity, and sex on eye gaze. *PLoS ONE 11*(12) Article 0168307. https://psycnet.apa.org/doi/10.1371/journal.pone.0168307

Wild, B., Rodden, F. A., Grodd, W., & Ruch, W. (2003). Neural correlates of laughter and humour. *Brain, 126*(10), 2121-2138. https://doi.org/10.1093/brain/awg226

Wingenbach, T. S. H., Ashwin, C., and Brosnan, M. (2018). Sex differences in facial emotion recognition across varying expression intensity levels from videos. *PLoS ONE 13*(1), Article e0190634. https://doi.org/10.1371/journal.pone.0190634

7. Appendices

7.1 Consentimento Informado

O objetivo deste estudo é compreender **como reconhecemos emoções em voz**, em expressões como **gargalhadas** e **choro**. Trata-se de um trabalho realizado no contexto de uma dissertação do Mestrado em Ciências em Emoções, a decorrer no ISCTE – Instituto Universitário de Lisboa, desenvolvida por Sofia Menezes (menezes.sofia@gmail.com), sob a orientação do Prof. Dr. César Lima (cesar.lima@iscte-iul.pt).

Ser-lhe-á pedido que preencha um questionário demográfico breve, seguido de uma tarefa em que ouvirá expressões vocais de emoção (gargalhadas e choro) e lhe pediremos que responda a perguntas simples sobre elas (e.g., se refletem uma emoção genuína ou não).

Ao todo, as tarefas duram 25-30 minutos.

Enquanto realiza as tarefas, pedimos que **use auscultadores/auriculares**, que se encontre num **local silencioso** e que **desligue as notificações sonoras** (por exemplo, email e mensagens de telemóvel).

Não existem riscos expectáveis associados à participação no estudo. Ainda que possa não beneficiar diretamente com a sua participação, as suas respostas contribuirão para o avanço do conhecimento científico sobre os processos psicológicos associados ao reconhecimento da autenticidade e valência das emoções.

Podem participar no estudo indivíduos cuja **primeira língua seja o português** e com **idade superior a 18 anos**.

A sua participação é voluntária e as suas respostas anónimas. Se escolher participar, pode interromper a participação em qualquer momento sem ter de prestar qualquer justificação.

Ao responder 'Sim' às seguintes afirmações ('Li e compreendi a informação prestada' e 'Aceito participar neste estudo') está a declarar que tem mais de 18 anos, que leu e compreendeu a informação apresentada nesta página, que considera que lhe foi dada a informação necessária sobre o estudo e que aceita participar voluntariamente no mesmo.

☐ **Li e compreendi a informação prestada**

☐ **Aceito participar neste estudo**