Prioritization of emotional signals by the human auditory system: evidence from a perceptual hysteresis protocol

Jean-Rémy Martin a,⁎, Guillaume Dezecache a,b,⁎⁎, Jérôme Dokić c, Julie Grèzes b

a Université Paris VI (UPMC), Institut d’Étude de la Cognition and Institut Jean-Nicod (ENS-EHESS-CNRS), Paris, France
b Laboratory of Cognitive Neuroscience (LNC) - INSERM U960 & IEC - Ecole Normale Supérieure (ENS), 75005, Paris, France
c Institut Jean-Nicod et École des Hautes Études en Sciences Sociales, Paris, France

Abstract

It is expected that natural selection has endowed our auditory apparatus with the ability to adaptively prioritize information that is crucial for survival and reproduction, such as vocal emotional signals emitted by our conspecifics, even in a noisy and dynamic natural environment (signals progressively emerge or fade away in noise as conspecifics move toward or away from us). Here, we tested the hypothesis that emotional signals are detected more easily (i.e., at lower signal-to-noise levels) and retained for a longer time (i.e., persisting in your sensory system at greater distance from the physical source) than signals bearing no emotional content, using a perceptual hysteresis protocol. Trials consisted of emotional signals (i.e., laughter and screams) or neutral signals (spectrally-rotated versions of the emotional stimuli) progressively emerging from white noise (ascending sequences) or progressively fading away in white noise (descending sequences). We demonstrated that vocal emotional signals were significantly detected at lower signal-to-noise levels than emotionally neutral signals in both ascending and descending sequences, suggesting that the human auditory system prioritizes signals bearing adaptive value.

1. Introduction

Emotional vocal signals are part of the behavioural repertoire of many mammal species (Hauser, 1997; Altenmüller, Schmidt, & Zimmermann, 2013), including humans (Sauter & Eimer, 2010; Sauter, Eisner, Ekman, & Scott, 2010; Scherer, Johnstone, & Klasmeyer, 2003), and are known to play a prominent role during social interactions. Among these vocalizations, screams and laughter, calls that signal distinct adaptive challenges in the environment (respectively, threat-related or agonistic contexts, and rewarding, social-bonding or playful interactions) seem to have been preserved throughout phylogeny (Morton, 1977; Vettin & Todt, 2005; Davila Ross, Owren, & Zimmermann, 2009). It is therefore expected that mechanisms favouring the production of attuned screams and laughter bouts (so as to produce signals that are effective in selectively affecting recipients) have co-evolved with mechanisms allowing for the accurate perception of these signals so that perceivers would be able to prioritize auditory information that is crucial for adaptation, particularly in noisy and dynamic environments, where signals are progressively emerging from or fading away in noise.

So far, research on the detection of vocal emotional signals (as well as visual emotional signals) has primarily focused on the detection of static stimuli (Scherer, Clark-Polner, & Mortillaro, 2011; Sacharin, Sander, & Scherer, 2012), that is to say, the detection of stimuli for which the amplitude does not evolve over time (Calvo & Estevés, 2005; Hock, Kelso, & Schöner, 1993). In addition, vocal emotions detection has seldom been studied under noisy conditions (Schuller, Seppi, Batliner, Maier, & Steidl, 2007; Tawari & Trivedi, 2010). More crucially, the detection of signals whose amplitude evolves over time in noise has received even less consideration. While studies addressing emotion detection with static stimuli mirror situations where the emitting source and the observer are fixed in space in a non-noisy environment, the present protocol mirrors more realistic or ecological situations where sources and observers are moving in a noisy environment: a source moving in a noisy space away from (or toward) an observer would lead to a progressive decrease (or increase) of signal-to-noise ratio (hereafter, SNR).

In this research, we tested the hypothesis that the auditory sensory system prioritizes vocal emotional signals of fear (scream) and amusement (laughter) when they are presented within a noisy dynamic setting. We predicted that emotional signals would be detected more easily (i.e., at lower signal-to-noise ratio) and maintained for longer (i.e., at lower signal-to-noise ratio too) than neutral signals. To test our predictions, we used an auditory protocol of hysteresis which provides a very effective way to test the
“persistence of a percept despite parameter change to values favouring the alternative pattern” (Hock et al., 1993, p. 63).

More precisely, hysteresis shows that the content of one’s perception at time t depends on the recent history of the perceptual system. In previous research, hysteresis effects have been shown to occur in many contexts: bistability (Gepshtein & Kubovy, 2005; Hock et al., 1993; Hock, Kogan, & Espinoza, 1997; Schwickert et al., 2012), form perception (Large, Aldcroft, & Vilis, 2005), letter recognition (Kleinschmidt, Büchel, Hutton, Friston, & Frackowiak, 1994) and facial emotions (Sacharin et al., 2012). Of particular interest here, Sacharin et al. (2012) showed that when subjects are presented with certain facial emotional expressions evolving over time from a particular emotion to another, they persist in perceiving the original emotion. For instance, when presented with faces evolving from the expression of anger to that of disgust, and from disgust to anger in a subsequent trial, the threshold at which subjects stop reporting seeing anger is lower in the anger-to-disgust trials, than the threshold at which they report starting perceiving anger in disgust-to-anger trials.

Hysteresis is usually investigated using designs comprising “ascending” and “descending” sequences, that is, sequences ordered in terms of a certain physical parameter. Here, ascending and descending sequences consisted of many steps with different signal-to-noise ratios between a target and a mask. The masks consisted of bursts of white noise of constant intensity. The targets were short emotional or neutral auditory signals. The SNR was progressively increased in ascending sequences or progressively decreased in descending sequences. A similar methodology has been used in previous experiments, notably in the visual domain (e.g., Kleinschmidt et al., 2002), and is known to efficiently produce hysteresis effects in subjects.

The use of this methodology revealed, in our experiment, significant greater detection at lower SNR-levels for emotional signals compared to neutral ones, in both ascending and descending sequences, suggesting that the human auditory system can prioritize signals bearing adaptive value.

2. Materials and methods

2.1. Participants

Eight participants (7 females, mean age of 25.25 years ± 0.619 SEM) participated in the study after having given their informed consent. None of our participants reported history of hearing problems, and all were naive regarding the purpose of the experiment. All participants were recruited from the database of the ‘Relais d’Information sur les Sciences de la Cognition’ (RISC, Paris, France). They received a compensation of 50€ for their participation.

2.2. Experimental setup

The experiment was conducted in a quiet experimental room. Stimuli were delivered by a MacBook Pro, processor 2.53GHz, Intel Core i5 equipped with a professional external sound card (One by APOGEE; A/D and D/A conversion, 44, 1/48 kHz, 24-bit) and presented through a pair of headphones (HD 250 linear II). All stimuli were displayed using Matlab (MathWorks Inc R2009b) with the Psychophysics toolbox extensions.

2.3. Stimuli

Emotional stimuli consisted of five fear stimuli (i.e., screams, mean of Fundamental frequency or F0 = 423.87 Hz; means of Formants F1-F4 = 925.59 Hz; 1644.72 Hz; 2634.93 Hz; 3469.31 Hz, respectively) and five amusement stimuli (i.e., laughter, mean F0 = 360.87 Hz; means F1-F4 = 963.42 Hz; 1905.27 Hz; 2749.84 Hz; 3895.22 Hz, respectively) produced by the same two women, and compiled by Sauter and colleagues (Sauter & Eimer, 2010) that we shortened to be of the same length (duration = 600 ms). From pilot studies, it appeared that 600 ms was a good trade-off between recognition of the emotion, and the length of trials (in pilot experiments, trials that were too long induced inattention in subjects). As shown below, stimuli were well categorized in their respective category and the length of a single trial did not exceed 30 seconds.

The neutral stimuli consisted of spectrally-rotated versions of the emotional stimuli: they were low-pass filtered at 3.8 kHz and were then spectrally rotated around 2 kHz (Blessen, 1972; Green, Rosen, Faulkner, & Paterson, 2013; Sauter & Eimer, 2010). Fig. 1 shows the spectrograms of a fear and laughter stimulus respectively as well as the spectrograms of their spectrally-rotated versions.

Spectral rotation transforms “the high-frequency energy to low-frequency energy and vice versa” around a specific frequency (Blessen, 1972, p. 5) so that spectra of rotated stimuli are mirror images of original stimuli spectra (see Fig. 1). From previous investigations (Green et al., 2013; Sauter & Eimer, 2010, Scott, Blank, Rosen, & Wise, 2000; Warren et al., 2006), it is acknowledged that spectral rotation preserves certain physical features present in original stimuli (i.e., duration, amplitude envelop and, perhaps more speculatively, pitch (Blessen, 1972)) but radically alters their global configuration, intelligibility and, crucially, their emotional significance (spectrally-rotated stimuli are perceived as being affectively neutral). This resulted in five neutral stimuli (henceforth, f-neutral, mean F0 = 289.60 Hz; means F1-F4 = 663.31 Hz; 1337.76 Hz; 2309.71 Hz; 2876.43 Hz, respectively) matched to the fear stimuli and five neutral stimuli matched to the amusement stimuli (henceforth, a-neutral, mean F0 = 303.02 Hz; means F1-F4 = 732.39 Hz; 1500.52 Hz; 2237.11 Hz; 2897.99 Hz, respectively).

2.4. Affective manipulation check

To make sure that neutral stimuli were actually perceived as affectively neutral and that emotional stimuli were well recognized as belonging to the fear or amusement category we conducted three pre-test experiments. In the first pre-test experiment, participants (n = 11; age range = 23–33 y.o.) were presented with the five fear and five affective neutral stimuli and the spectrally-rotated versions of these stimuli in a random order and asked, in a forced-choice task, to categorize each stimulus in one of the three following categories: fear, amusement or neutral. Participants categorized stimuli above chance-level (t0 = 5.724, p < .001 for fear; t10 = 13.895, p < .001 for f-neutral; t0 = 9.682 p < .001 for amusement; t0 = 5.073; p < .001, for a-neutral). In the second pre-test experiment, to ensure that spectrally-rotated stimuli could not convey any fear or amusement tone, another group of participants (n = 10; age range = 22–34 y.o.) was presented only with the spectrally-rotated stimuli; they were asked to classify each stimulus in one of two categories, i.e., fear or amusement, in a forced-choice task. We found that participants classify stimuli at chance-level (t0 = 1.097, p > .1 for f-neutral; t0 = 0.178, p > .1 for a-neutral), confirming that spectrally-rotated stimuli were perceived as affectively neutral. Finally, we found that, in 10 other participants (age range = 20–50 y.o.), emotional stimuli were perceived as more negatively (for fear stimuli) and more positively (for amusement stimuli) arousing than their spectrally-rotated counterparts (fear stimuli: t0 = −2.311, p < .05; amusement stimulus: t0 = 4.595, p < .001).

In addition (as a control condition), and to further confirm that any differential effects in hysteresis and detection thresholds that we might find between emotional signals and spectrally-rotated stimuli (see below) are indeed due to the emotional content of emotional stimuli, rather than to their human character (given that spectrally-rotated stimuli might be perceived as “robotic sounds”), the same subjects performed the control condition session investigating hysteresis levels
for additional neutral stimuli. These stimuli consisted of auditory letters (E, U, O, X, K, T, Z, B, H, mean F0 = 210.55; means F1-F4 = 642.16; 1974.81; 2971.76; 4199.21, respectively) that were recorded in a soundproof room by a female student. The use of letters as additional control stimuli also has the advantage of ruling out the possibility that potential differences of detection between emotional stimuli and their spectrally-rotated versions could be due to the former being more “meaningful” than the latter. Indeed, letters are meaningful elements.

All stimuli were displayed using Matlab (MathWorks Inc R2009b) with the psychophysics toolbox extensions.

2.5. Procedure

To design our hysteresis protocol, we used the modified method of limits (Hock & Schöner, 2010), which contrasts with the traditional method of limits (Fechner, 1860). While the latter method also produces hysteresis-like phenomena, the difference in detection thresholds between ascending and descending sequences that characterizes hysteresis has been attributed to a number of artefacts (Woodworth & Schlosberg, 1954), including (1) perseveration in response, and (2) inference production from trial duration. The first bias results from the fact that ascending and descending sequences were not randomly displayed in the traditional methods of limits. Therefore, the persistence effect might have simply reflected persistence or habituation in the response of subjects. The second bias is a consequence of the unequal duration of (ascending as well as descending) sequences. In other words, each value to reach had a specific number of steps. In this context, the responses of subjects could be essentially based on the duration of the current trial, longer sequences (for instance) being more likely to reach values where the percept has changed. The modified method of limits avoids the two biases by, respectively, randomizing the presentation of ascending and descending sequences and by equalizing the duration of all sequences (the progressive increasing or decreasing of values beginning at different times within the sequences according to the final value to reach). Finally, and importantly, in the modified method of limits, contrary to the traditional method, participants...
Table 1
Sequences of SNR-levels (in dB) for the different conditions.

<table>
<thead>
<tr>
<th>Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascending sequences (SNR-values)</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
<td>−30</td>
</tr>
<tr>
<td>Descending sequences (SNR-values)</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
<td>−14</td>
</tr>
</tbody>
</table>

Table shows the series of SNR-levels from stimulus 1 to 16 (stimulus 16 was the test stimulus for the detection judgments) across the different conditions. All Ascending sequences started with a SNR-value of −30 and all Descending sequences started with a SNR-value of −14. The grey region in each row indicates the beginning of increment (in ascending sequences) or decrement (in descending sequences) of the SNR.

 respond at the end of each trial and not during the current sequence of ascending or descending steps avoiding specific decision or response-time biases (Hock & Schöner, 2010). In sum, the modified method of limits has been specifically designed to show authentic persistence or hysteresis effects, that is, persistence effects not attributable to certain artefacts but, instead, indicating the presence of specific processes (Hock et al., 1993; Hock & Schöner, 2010).

In the present study, participants were presented with sequences composed of 16 steps (inter-stimulus interval (ISI) = 600 ms). A step consisted of a specific (neutral or emotional) stimulus embedded in a burst of white noise (60 dB SPL).

Two main types of condition or types of trial were tested. In ascending sequences, the SNR measured in dB was progressively increased by steps of 1 dB. In descending sequences, the SNR was progressively decreased by steps of 1 dB too. At the end of the ascending and descending sequences, participants had to make a judgment about the presence or absence of the signal in noise for the last 16th step.

The progressive increase or decrease of the SNR in a sequence began at a specific step within a trial according to the final value of intensity the stimulus should reach. Nine SNR-values were tested (−30, −25, −24, −23, −22, −20, −19, −18, −16) and repeated 20 times (10 times within ascending sequences and 10 times within descending sequences) for each kind of stimulus (fear, f-neutral, amusement, a-neutral). The choice of the values was encouraged by unpublished pilot data.

Ascending and descending sequences as well as the different SNR values were randomly presented. Table 1 sums up the characteristics of the different ascending and descending sequences.

The experiment was composed of two hysteresis blocks (block A and B) repeated twice (= four sub-blocks). Together with the control condition session, the order of blocks was counterbalanced across participants. Block A was composed of fear and f-neutral stimuli presented in two sub-blocks; block B of amusement and a-neutral stimuli in two other sub-blocks. In other words, each sub-block was composed of ascending and descending sequences; blocks and sub-blocks differed only in terms of the types of stimuli they displayed. Participants did not perform more than one block per day. Four days were thus necessary to complete the entire experiment (total duration = around 5 hours). The control condition session followed the exact same procedure.

2.6. Statistical analysis

To explore differences in perceptual persistence across conditions, we fitted detection curves upon sigmoid functions so as to determine, for each participant and for each condition (fear, f-neutral, amusement, a-neutral), SNR-values where detection rates cross 75% (75% of detection being the conventional threshold used in psychophysics, see Kingdom & Prins, 2010). Before doing so, we made sure that sigmoid functions provided a
reasonable fit of our participants’ data (mean $R^2 = 0.657$; range = 0.406–0.923). Then, SNR-values were submitted to an ANOVA with blocks (block fear/f-neutral, block amusement/a-neutral), Content (emotion, neutral) and Direction (ascending, descending) as within-participants factors. Bonferroni corrections were also employed to account for multiple testing. Post-hoc comparisons (two-tailed t-tests) were performed for the analysis of 3-way interactions.

3. Results

Fig. 2 shows the SNR-value for which the different categories of stimuli reach (ascending sequences) and drop below (descending sequences) 75% of detection. We directly analysed differences between emotions and neutral stimuli in terms of detection (at which thresholds stimuli reach 75% of detection for ascending sequences) and persistence (at which thresholds stimuli drop below 75% of detection for descending sequences) (see Fig. 2). The ANOVA revealed main effects of Direction ($F_{(1,7)} = 34.728, p = .001$) (stimuli show 75% of detection rate at lower thresholds in descending sequences), Blocks ($F_{(1,7)} = 18.728, p = .003$) (stimuli show 75% of detection rate at lower thresholds in block fear/f-neutral), and Content ($F_{(1,7)} = 127.061, p < .001$) (stimuli show 75% of detection at lower thresholds for emotional stimuli). We also found a double interaction between factors Block and Content ($F_{(7,49)} = 22.178, p = .002$).

Of main interest here, statistical analysis showed a triple interaction between Block, Content and Direction factors ($F_{(1,7)} = 12.396, p = .01$). Post-hoc tests revealed that, in ascending sequences, detection of fear stimuli reaches 75% of detection at lower SNR-levels compared to f-neutral stimuli ($t_{(7)} = 5.274, p = .001$); similarly, detection of amusement stimuli reaches 75% of detection at lower SNR-levels compared to a-neutral stimuli ($t_{(7)} = 14.413, p < .001$). In descending sequences, fear stimuli dropped below 75% of detection at lower SNR-levels compared to f-neutral stimuli ($t_{(7)} = 5.273, p < .05$), and amusement stimuli dropped below 75% of detection at lower SNR-levels compared to a-neutral stimuli ($t_{(7)} = 6.065, p = .001$).

More specifically, in ascending sequences fear stimuli reached 75% of detection from the simulated SNR-value $-23.29 (+/-0.38$ SEM); f-neutral stimuli from $-21.55 (+/-0.59$ SEM); amusement stimuli from $-22.59 (+/-0.81$ SEM); a-neutral stimuli from $-16.47 (+/-0.82$ SEM). In descending sequences, fear stimuli dropped below 75% of detection from the SNR-value $-27.70 (+/-1.33$ SEM); f-neutral stimuli from $-25.27 (+/-0.81$ SEM); amusement stimuli from $-26.26 (+/-0.64$ SEM); a-neutral stimuli from $-22.60 (+/-0.75$ SEM).

One might argue that the effects we found were due to the non-human character of the spectrally-rotated stimuli, and that the effects we found for emotional stimuli could have simply been obtained using stimuli that carry any sort of human information. To investigate this possibility, we compared estimated SNR-values obtained for emotional stimuli with those obtained when participants were presented with another type of neutral stimuli, namely auditory letters. We found that, for ascending sequences, 75% of detection was reached at lower SNR-values for fear stimuli compared to letters ($t_{(7)} = -6.953, p < .001$), and for amusement stimuli compared to letters ($t_{(7)} = -4.038, p = .005$). In the case of descending sequences, detection dropped below 75% at lower SNR-values for fear stimuli compared to letters ($t_{(7)} = 3.157, p < .05$), and for amusement stimuli compared to letters ($t_{(7)} = 3.788, p < .01$). More specifically, in ascending sequences, letters reached 75% of detection from the SNR-value $-20.14 (+/-0.34$ SEM) and, in descending sequences, letters dropped below 75% of detection from the SNR-value $-23.14 (+/-0.68$ SEM) (see Fig. 2).

4. Discussion

By designing an experiment in which the amplitude of specific signals evolved monotonically over time in noise, we revealed that emotional signals were detected more easily than neutral signals and maintained for longer over noise compared to neutral signals. Indeed, they were detected at lower signal-to-noise ratios in both ascending and descending sequences compared to neutral stimuli. The present findings support the idea that, within a noisy dynamic setting, the auditory system prioritizes evolutionarily and ecologically relevant information such as emotional information. Emotional signals possess an ecological value that neutral signals do not exhibit. Vocal emotional stimuli expressing fear are crucial signals for survival: for instance, a scream is typically produced in response to a threat within the close environment (Sauter et al., 2010). Similarly, vocalizations expressing amusement can signal social bonding opportunities (Mehu & Dunbar, 2008a, 2008b; Dunbar et al., 2012) or playful interactions (Vettin & Todt, 2005; Davilla-Ross, Alcock, Thomas, & Bard, 2011). Consequently, being equipped with a perceptual system capable of detecting such emotional stimuli at low SNR-levels and to maintain them at low-SNR levels over a noisy environment is likely to be highly beneficial in terms of fitness consequences.

The better detection of emotional stimuli compared to neutral stimuli could result in an enhancement of perceptual processing in presence of emotional stimuli, rather than to a bias favouring the detection of emotional stimuli over neutral stimuli. This is consistent with evidence showing that emotions trigger enhanced perceptual processing (Calvo & Estes, 2005; Zeelenberg, Wagenmakers, & Rotteveel, 2006). However, in using perceptual hysteresis protocols enabling the measure of d-primes and confidence, future studies should explore whether the higher detection rate for emotional signals actually results from an enhanced perceptual sensitivity in presence of emotional signals (and, as a consequence, in increased confidence judgments), or from the fact that judgments of confidence themselves are modulated (with higher proportion of false positives or false alarms for emotional than for neutral signals). If such prioritization of emotional signals was merely a matter of decision rather than that of an enhanced perceptual sensitivity, it would nonetheless be a very interesting finding. In particular, when the signal is disappearing (in descending sequences) or appearing (in ascending sequences), the production of more false positives for emotional signals than for neutral signals would be actually ‘ advantageous’ in avoiding misses that would be evidently more disadvantageous when signals convey crucial information. In short, emotional signals would make subjects more liberal than neutral signals do.

A potential shortcoming of the present study is that the low-pass filter of 3.8 kHz could eliminate some of the acoustic information used in voice perception, such as the upper formant frequencies (F4+) making the neutral stimuli relatively less discriminable than the emotional signals. However, vocal letters do possess these upper frequencies (a spectral analysis of letters revealed that they had even more high-frequency components than emotional signals). Therefore, our results cannot be accounted for by the fact that neutral stimuli were less discriminable than emotional stimuli.

One could also argue that our effects are driven solely by neutral stimuli (in particular, the spectrally-rotated counterparts of the emotional stimuli) presenting slightly lower pitch for some formants than emotional vocalizations (see Sections 2.3 and 2.4). If this might account for an overall greater discriminability of these emotional signals, this likely does not explain the better detection of emotional signals in our experiment. Indeed, higher formants are also found to be lower in fear and amusement stimuli compared to letters. Letters, however, do not trigger comparable perceptual persistence, further suggesting that certain high-level properties (the recognition that a certain emotional signal is being produced) play a role in perceptual persistence. Future studies should however investigate in more detail the impact and the contribution of the various acoustic parameters in perceptual persistence. Such studies should bring insight into certain relevant acoustical features that would help make certain signals
particularly suitable for the conveyance of adaptive information (Rendall, Owren, & Ryan, 2009; Bryant, 2013).

In addition, it could also be argued that our second type of neutral stimuli, i.e., letters, present participants with relatively little semantic complexity compared to emotional vocal signals. Also, they were of shorter duration than the emotional vocalizations. If further study should eliminate these potential experimental biases, we shall argue that other methodological strategies might bring other potential biases. As for the issue of semantic complexity, it is difficult to find purely neutral stimuli. Words can, for instance, trigger affective reactions in certain participants (because of, e.g., personal history) without the possibility to properly control for it. Additionally, it is unlikely that the length of the stimuli plays an important role in shaping persistence effects: as far as letters were fully heard and understood, their lesser perceptual persistence cannot be accounted by a simple limitation in auditory processing.

Another limitation of this study is the over-representation of females in our sample of participants (7 females vs. 1 male only). Women are known to be more responsive that men to emotional stimuli (Hall, 1978, 1990; Barrett, Lane, Schreest, & Schwartz, 2000; Bradley, Codispoti, Sabatinielli, & Lang, 2001). This could have induced stronger persistence biases in our results which would not then be representative of a mixed human population. For example, men could perceive fear stimuli as less arousing (and therefore show less persistence). Conversely, laughter bouts produced by the opposite sex could signal reproductive opportunities (Grammer, 1990). Again, they could have been perceived as conveying an increased level of emotion (if they were compared with the same vocalizations when produced by men), and this could also have an effect on persistence of emotional stimuli compared to that of neutral stimuli. In addition, there is evidence that females and males differ in auditory processing of sounds. For example, females typically have better hearing sensitivity and differ from males in perception of sounds embedded in noise (see (McFadden, 1998) for review). In any case, further work is needed to measure the extent to which the over-representation of females in our participant sample and in our stimuli influenced the greater persistence of emotional content.

Finally, future research should try to replicate the present findings with different procedures than the modified method of limits. This method has been mainly designed in the context of bistable stimuli (Hock et al., 1993, 1997; Hock & Schöner, 2010), but we faithfully adapted it in a context of detection. We preferred the modified method of limits to a staircase procedure for the following reasons: first, the modified method of limits has been specifically designed in the context of perceptual hysteresis; second, staircase procedures could have suppressed hysteresis effects. Of course, we could adapt a staircase procedure so that we have progressively decreasing or increasing SRN-levels within trials of identical length, but here comes the third reason. In such a staircase procedure, ascending and descending sequences should be blocked rather than randomly displayed trial per trial, potentially leading to the habituation bias described in the procedure section and/or to intertrial’s hysteresis effects. If we modify the procedure so that ascending and descending sequences are randomly displayed trial per trial the procedure, in fact, becomes very close to the modified method of limits. However, this is ultimately a question that has to be explored empirically.

Acknowledgments

The authors would like to thank Daniel Pressnitzer for technical assistance and for important comments about the manuscript. Disa Sauter for sharing the stimuli, Valentin Wyatt for analytics tools, Victor Benichoux, Hadrien Orvoën, Elisabeth Pacherie, Ariadna Fernandez for helpful comments. The research was supported by an ANR-11-0001-02 PSL*, an ANR-10-LABX-0087, an ED3C/UPMC scholarship (J.R.M.) and a DGA-MRIS scholarship (C.D.).

References


Please cite this article as: Martin, J-R., et al., Prioritization of emotional signals by the human auditory system: evidence from a perceptual hysteresis protocol, Evolution and Human Behavior (2014), http://dx.doi.org/10.1016/j.evolhumbehav.2014.07.005


