

Involvement of the Serotonin Transporter Gene in Accurate Subcortical Speech Encoding

Lenka Selinger,^{1,2} Katarzyna Zarnowiec,^{1,2}  Marc Via,^{1,2,3}  Immaculada C. Clemente,^{1,2,3} and  Carles Escera^{1,2,3}

¹Institute of Neurosciences, University of Barcelona, 08035 Barcelona, Spain, ²Brainlab, Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, 08035 Barcelona, Spain, and ³Institut de Recerca Sant Joan de Déu, 08950 Esplugues de Llobregat, Spain

A flourishing line of evidence has highlighted the encoding of speech sounds in the subcortical auditory system as being shaped by acoustic, linguistic, and musical experience and training. And while the heritability of auditory speech as well as nonspeech processing has been suggested, the genetic determinants of subcortical speech processing have not yet been uncovered. Here, we postulated that the serotonin transporter-linked polymorphic region (5-HTTLPR), a common functional polymorphism located in the promoter region of the serotonin transporter gene (SLC6A4), is implicated in speech encoding in the human subcortical auditory pathway. Serotonin has been shown as essential for modulating the brain response to sound both cortically and subcortically, yet the genetic factors regulating this modulation regarding speech sounds have not been disclosed. We recorded the frequency following response, a biomarker of the neural tracking of speech sounds in the subcortical auditory pathway, and cortical evoked potentials in 58 participants elicited to the syllable /ba/, which was presented >2000 times. Participants with low serotonin transporter expression had higher signal-to-noise ratios as well as a higher pitch strength representation of the periodic part of the syllable than participants with medium to high expression, possibly by tuning synaptic activity to the stimulus features and hence a more efficient suppression of noise. These results imply the 5-HTTLPR in subcortical auditory speech encoding and add an important, genetically determined layer to the factors shaping the human subcortical response to speech sounds.

Key words: 5-HTTLPR; frequency-following response; genetics; individual differences; SLC6A4; speech perception

Significance Statement

The accurate encoding of speech sounds in the subcortical auditory nervous system is of paramount relevance for human communication, and it has been shown to be altered in different disorders of speech and auditory processing. Importantly, this encoding is plastic and can therefore be enhanced by language and music experience. Whether genetic factors play a role in speech encoding at the subcortical level remains unresolved. Here we show that a common polymorphism in the serotonin transporter gene relates to an accurate and robust neural tracking of speech stimuli in the subcortical auditory pathway. This indicates that serotonin transporter expression, eventually in combination with other polymorphisms, delimits the extent to which lifetime experience shapes the subcortical encoding of speech.

Introduction

Why do some people have a better ear for music and some learn to pronounce a new language so much better than others? Clearly,

Received May 17, 2016; revised Aug. 24, 2016; accepted Aug. 27, 2016.

Author contributions: L.S., K.Z., I.C.C., and C.E. designed research; L.S., K.Z., and M.V. performed research; L.S. and K.Z. analyzed data; L.S. and C.E. wrote the paper.

This work was supported by grants from the Spanish Ministry of Economy and Knowledge (Grant PSI2015-63664-P), the Catalan Government (Grant SGR2014-177), and the ICREA Acadèmia Distinguished Professorship awarded to C.E.

The authors declare no competing financial interests.

Correspondence should be addressed to Carles Escera, Brainlab, Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Passeig de la Vall d'Hebron 171, 08035 Barcelona, Catalonia, Spain. E-mail: cescera@ub.edu.

DOI:10.1523/JNEUROSCI.1595-16.2016

Copyright © 2016 the authors 0270-6474/16/3610782-09\$15.00/0

individuals differ in their musical ability, and in their capacity to learn a new language and to correctly pronounce it (Díaz et al., 2008). More generally, it could be said that people differ in their capacity for sound and, importantly, speech perception (Pallier et al., 1997). An accurate and fine-grained neural representation of the sound signal is the basis for speech perception.

The subcortical auditory pathway has been shown to encode acoustic features of speech and nonspeech stimuli with great accuracy (Skoee and Kraus, 2010a). Even in the typically developing population, however, the strength with which features important for language perception are represented varies systematically with language and reading competence (Banai et al., 2009; Chandrasekaran et al., 2009; Hornickel et al., 2009, 2011). Moreover, the auditory brainstem response, comprising both the transient

and the periodic frequency following response (FFR) to sound, has been shown to be altered in children with specific language impairment (Basu et al., 2010), attention deficit disorder with hyperactivity (Jafari et al., 2015), and auditory processing disorder (Rocha-Muniz et al., 2012). In preschooler children, it has been proposed as a biomarker to predict literacy (White-Schwoch et al., 2015).

The FFR is a neural response to periodic sounds that can be recorded from the human scalp (Chandrasekaran and Kraus, 2010), and whose generators have been attributed to auditory brainstem nuclei, particularly the inferior colliculus (IC; Smith et al., 1975; Møller et al. 1988), yet a cortical contribution cannot be disregarded for certain stimulus characteristics (Coffey et al., 2016). It faithfully reflects the envelope and frequency contents of sounds. The FFR grants a noninvasive, objective means for examining the neural encoding of sound at subcortical stages of the auditory pathway. These nuclei play an important role as a computational hub, receiving bottom-up auditory input from brainstem structures, as well as direct corticofugal feedback projections from the auditory cortex. Remarkably, the fidelity with which the ascending auditory pathway represents the physical characteristics of sounds in the FFR is experience dependent, as musical training (Parbery-Clark et al., 2013), bilingualism (Krizman et al., 2012), the socioeconomic status of the person (Skoe et al., 2013a), and the context in which a stimulus is presented (Skoe and Kraus, 2010b) shape the subcortical response to sound. It has been hypothesized that increased variability in the FFR leads to a compromised interaction between sensory and cognitive processing, resulting in disparities in cognitive abilities between individuals (Skoe et al., 2013a). Whether genetic factors contribute to individual differences regarding sound processing at a subcortical level, however, has not been established so far.

It has been shown that the IC receives major serotonergic innervation originating in the raphe nuclei (Klepper and Herbert, 1991; Hurley and Pollak, 1999). This suggests that serotonin [5-hydroxytryptamine (5-HT)] is crucial for the modulation of information processing in the ascending auditory pathway (Obara et al., 2014). In particular, animal studies have demonstrated that serotonin alters multiple aspects of the neural response to sound. Not only does it change the number of spikes evoked by auditory stimuli (Hurley and Pollak, 1999, 2001), but also the latency and precision of initial spikes and the timing of spike trains and therefore all response properties contributing to the encoding of sensory stimuli (Hurley and Pollak, 2005). As the serotonin transporter [5-HT transporter (5-HTT)] is considered to be crucial for the regulation of serotonin availability and a common functional polymorphism has been identified in the serotonin transporter gene (SLC6A4; Heils et al., 1996), we hypothesized that it might be implicated in subcortical auditory processing.

The aim of this study was to determine whether the 5-HTT linked polymorphic region (5-HTTLPR) is involved in speech encoding at subcortical stages. Two measures of encoding accuracy and robustness of neural phase locking to the stimulus periodicity, namely, the signal-to-noise ratio (SNR) and pitch strength, revealed a more finely tuned auditory processing at subcortical stages for individuals with low serotonin transporter expression compared with those with medium to high expression.

Materials and Methods

Participants

Seventy-nine young adults ranging in age from 18 to 31 years [mean (\pm SEM) age, 22.32 \pm 3.28 years; 59 females] participated in the study. They were mostly psychology students, recruited from a larger sample of volunteers who were interviewed according to an adapted version of the

Clinical Interview of the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, for the exclusion of participants with neurological or psychiatric illness or drug consumption. Participants were compensated economically for their time (\sim 7€/h). All participants gave their written informed consent at each phase of the study (interview, buccal cell extraction, and EEG recordings), according to the Declaration of Helsinki and approved by the Bioethics Committee of the University of Barcelona. Before the recording session, participants completed the State and Trait Anxiety inventory (Spielberger et al., 1988) and a questionnaire regarding musical experience. All participants had normal or corrected-to-normal vision and normal audition as assessed at the beginning of the recording session with a standard pure-tone audiometry using DT48-A headphones (Beyerdynamic). After exclusion by diagnostic criteria and after obtaining the 5-HTTLPR polymorphism, participants were invited for an EEG recording session. A total of 21 participants were excluded, as follows: 5-HTTLPR testing was inconclusive for 14 participants; 5 participants had very low SNRs (\leq 2); and 2 more participants had other technical issues regarding their EEG recordings.

Genotyping

Samples were collected using buccal cell cheek swabs (BuccalAmp DNA Extraction Kit, Epicentre), and DNA was extracted following manufacturer specifications. DNA samples were sent to an external genomics core facility (Progenika Biopharma) to determine genetic variants.

The 5-HTTLPR is a variable repeat sequence in the promoter region of the SLC6A4 gene. There are two allelic variants of this region: a short (S) variant comprising 14 copies of a 20–23 bp irregular repeat unit, and a long (L) variant comprising 16 copies. The 5-HTT removes 5-HT released into the synaptic cleft. Being a carrier of either the S or the L allele results in differential 5-HTT expression (Heils et al., 1996). Presence of the S allele is associated with reduced transcriptional activity and 5-HTT expression, relative to the L allele (Lesch et al., 1996), and therefore an increased availability of 5-HT in the synaptic cleft. Additionally, the L allele contains an A/G single nucleotide polymorphism (SNP; rs25531) which may render the L_G allele functionally similar to the S allele (Hu et al., 2005).

Both the 5-HTTLPR (rs63749047) and the A/G SNP (rs25531) were determined through PCR amplification, digestion with the restriction enzyme MspI, and migration in an agarose gel. Digestion was performed via incubation of 10 μ l of the PCR product with 5 U of MspI (New England Biolabs) for 1.5 h at 37°C. The PCR products and digestion were run in 2.5% agarose gels and separated through electrophoresis. Visualization took place in a UV transilluminator through SYBR Safe staining.

Thus, genotypes resulting from either the L or S allele and the A/G SNP contained in the L allele could be one of the following six: L_A/L_A, L_A/L_G, L_G/L_G, L_A/S, L_G/S, and S/S (Hu et al., 2005). The participants were initially assigned to one of three groups according to their level of 5-HTT expression, as follows: low (S/S, L_G/S, L_G/L_G; 23 participants); medium (L_A/L_G, L_A/S; 21 participants); and high (L_A/L_A; 14 participants). As pairwise comparisons did not yield any statistically significant differences regarding SNR ($p = 0.93$) and autocorrelation ($p = 0.86$) for the medium-expressing and the high-expressing groups, these participants were pooled together in one group (medium–high-expressing group, 35 participants).

Stimuli

Brainstem responses were recorded in response to the consonant-vowel syllable /ba/, which was generated with the Klatt speech synthesizer (Klatt, 1980) and designed as in the study by Slabu et al. (2012; Fig. 1A). The stimulus had a duration of 170 ms and a fundamental frequency (F_0) of 100 Hz. The third (F3), fourth (F4), and fifth (F5) formants were set to 2900, 3500, and 4900 Hz, respectively. To elicit a large onset response, the first 5 ms included a rapid glide in the first (F1; from 400 to 1700 Hz) and second (F2; from 1700 to 1240 Hz) formant. As sounds never occur in isolation, for a more natural listening situation, and as differences between participants in signal extraction accuracy have been shown especially during challenging listening conditions (Song et al., 2011), a looped semantically anomalous speech babble was played in the background, 10 dB SPL softer than the syllable. The babble was recorded in a sound-

attenuated booth, at a 44 kHz sampling rate and 16 bit accuracy using Audacity Software. It consisted in a short, semantically anomalous text, which was read in a natural and conversational manner by six native Spanish speakers (two male, four female), and then circularly looped with each speaker starting 10 s after the previous one and normalized off-line by root mean square (rms) amplitude normalization in Matlab version 7.4 (MathWorks; RRID: SCR_001622). To avoid any kind of interaction with the response to the syllable, the babble was looped with no silent intervals and the stimulation with the syllable started at a random phase of the babble.

The syllable was presented binaurally, with alternating polarities via ER-3A ABR insert earphones (Etymotic Research) at an intensity of 80 dB SPL and in the following two different timing conditions: constant and jittered stimulus onset asynchrony (SOA). In the constant condition, the SOA was set to 300 ms, while it ranged from 256 to 344 ms in the jittered condition (mean SOA, 300 ms). On the whole, 1008 trials/condition were presented in two separate blocks (the order of blocks was randomized across participants). Each block lasted ~5 min, and participants were allowed to rest between blocks. During the whole recording session, participants sat in a comfortable chair, in an acoustically and electrically shielded room. As brainstem responses to sound are of a preattentive nature, participants were instructed to relax and focus on a silent movie with subtitles, while ignoring the auditory stimulation.

EEG recordings

The FFRs and cortical responses were extracted from continuous EEG recordings, acquired with a SynAmpsRT amplifier (Compumedics Neuroscan) and Scan version 4.4 software (Compumedics Neuroscan). Data were recorded from 36 scalp Ag/AgCl electrodes mounted in a nylon cap (Quick-Cap, Compumedics Neuroscan) at the standard 10–20 system locations. Two additional electrodes were positioned at the left and the right mastoids (M1 and M2, respectively). The electrooculogram (EOG) was measured with two bipolar electrodes placed above and below the left eye, and two horizontal electrodes placed on the outer canthi of the eyes. The ground electrode was located between Fz and FPz, and the right earlobe (A2) served as the on-line reference. All impedances were kept to <10 k Ω during the whole recording session, and data were on-line bandpass filtered from 0.05 to 3000 Hz and digitized with a sampling rate of 20 kHz.

Data processing and analyses

FFRs. Data analysis was performed off-line using EEGLab version 7 toolbox (Delorme and Makeig, 2004; RRID: SCR_007292) running under Matlab version 2012a. The continuous recordings extracted from the CPz electrode were filtered off-line with a bandpass Kaiser window finite impulse response filter from 70 to 1500 Hz and epoched from 40 ms before the stimulus onset to 180 ms after stimulus onset.

An independent component analysis second-order blind identification (SOBI; Delorme and Makeig, 2004; Delorme et al., 2007) procedure was implemented to clean the recorded EEG from artifacts and to retain the signal activity corresponding to the FFR only. This was applied for each participant separately, on demeaned and detrended epochs of 10–170 ms from sound onset. All of the EEG and EOG channels were used in this step, and a total of 1008 epochs per condition were created. All

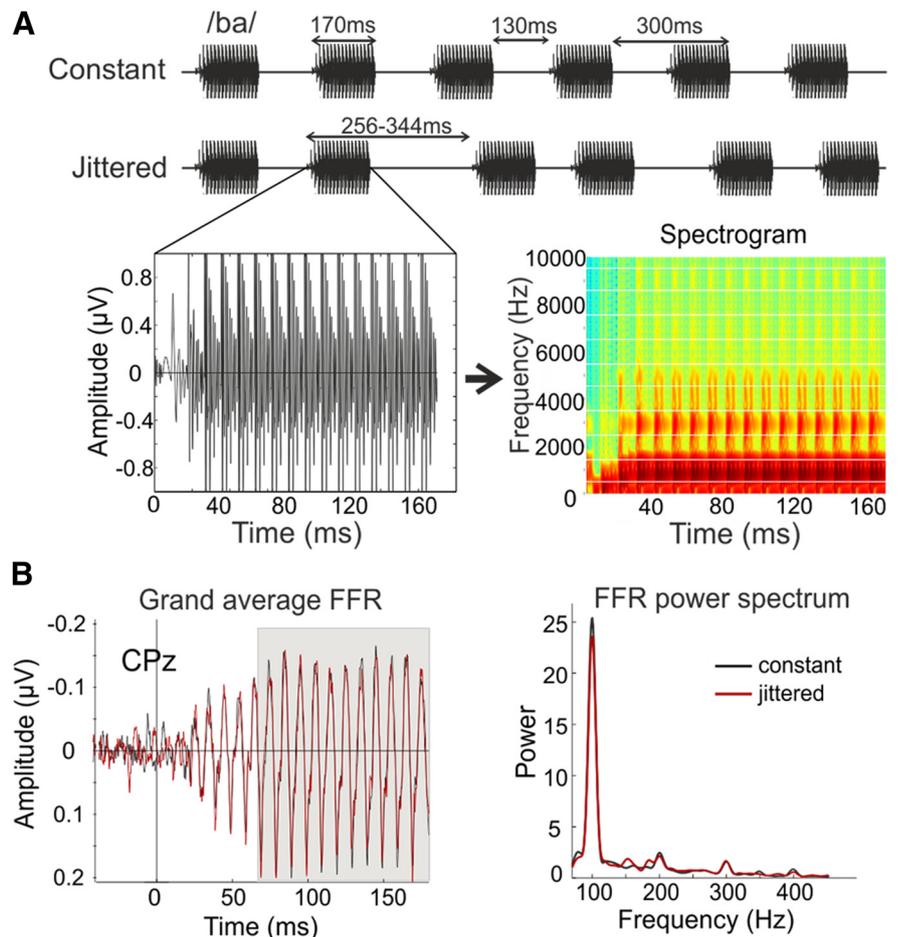


Figure 1. Experimental design and FFR results. **A**, Design. Participants were presented with the consonant-vowel syllable /ba/ with a duration of 170 ms. The fundamental frequency was 100 Hz, while the third, fourth, and fifth formants were set to 2900, 3500, and 4900 Hz respectively. The syllable was presented at 80 dB SPL over a 10 dB SPL lower looped semantically anomalous speech babble in either a constant (SOA, 300 ms) or a jittered (SOA, 256–344 ms; mean SOA, 300 ms) condition. **B**, FFR grand average. Left, FFR responses for both conditions, measured at the CPz electrode. The steady-state part on which we focused in our analyses is highlighted in gray. Right, Power spectrum of the FFR for both timing conditions with slightly, but not significantly higher power in the constant condition.

epochs possibly containing artifacts were removed from the dataset via the automatic rejection algorithm implemented in EEGLAB (i.e., epochs with amplitudes exceeding $\pm 95 \mu\text{V}$), and epochs where the peak-to-peak values were exceeding $150 \mu\text{V}$ were removed from each participant's data (up to a maximum of 30% of all epochs). The cleaned data were input into the SOBI algorithm, and 40 independent components were found for each subject. Component activations were averaged, windowed with a Hanning taper, and zero padded to obtain a resolution of 1 Hz/bin, then fast Fourier transform (FFT) was computed on each of the components. First, the signal was estimated by reading out mean power values corresponding to a 10 Hz window around the fundamental frequency (100 Hz). Then an estimation of the baseline (noise) was performed, by taking the mean power value of a 10 Hz window on each side of the selected peak (with a separation of 20 Hz from the peak). Next, the SNR was calculated to select only components in which the amplitude of the signal at the frequency of interest was larger than the amplitude of the noise (according to an *F* test; significance level $p < 0.05$). All remaining components were treated as irrelevant for FFR analysis and pruned. On average, 36 components were retained for each subject. Finally, weights corresponding to FFR components were applied to the initially created epochs. The epochs were baseline corrected to a 40 ms interval preceding the sound onset (Russo et al., 2008).

Responses to stimuli of alternating polarities were averaged together to minimize stimulus artifacts and cochlear microphonic, preserving the FFR to the stimulus envelope (Aiken and Picton, 2008). In the FFR to the

syllable, two parts can be distinguished: the consonant–vowel transition (10–65 ms), which poses a particular perceptual challenge to the listener (Assmann and Summerfield, 2004); and the steady-state part corresponding to the processing of the vowel (65–179 ms), which is characterized by constant temporal features (Song et al., 2011).

Epochs for the constant and the random blocks were averaged separately, and the grand average was calculated. To examine the FFR in the frequency domain, FFT (Cooley and Tukey, 1964) was applied to demeaned, zero-padded (1 Hz resolution) averages, windowed with a Hanning taper. The mean response amplitude was computed using a 10 Hz window surrounding F_0 (90–110 Hz) and the subsequent three harmonics: H2, H3, and H4. Higher harmonic components were not reliably present in all subjects; therefore, only the response to F_0 was statistically analyzed. As an index for response robustness, the SNR was determined by calculating the mean value of noise in a window of 10 Hz on either side of the peak, with a separation of 20 Hz between the window surrounding the peak and the window taken for noise, and subsequently dividing the mean signal value (power; e.g., squared amplitude) by the mean noise value. Overall condition and genetic effects were assessed by means of a two-way repeated-measures ANOVA with the within-subjects factor timing (constant vs jittered) and the between-subjects factor 5-HTT expression (low vs medium–high). Years of musical training were introduced to the ANOVA as a covariate. As the experimental manipulation of presentation timing of the syllable did not yield any significant results (SNR: $F_{(1,65)} = 0.04$, $p = 0.83$; pitch strength: $F_{(1,65)} = 1.3$, $p = 0.26$), we merged these two conditions. To exclude the possibility that differences in the prestimulus activity between groups could account for any measured effect, we calculated the rms values for both groups (mean \pm SEM: low-expression group, 0.232 ± 0.027 ; medium–high-expression group, 0.267 ± 0.022) and found no significant difference ($F_{(1,65)} = 0.714$; $p = 0.493$).

Pitch is an important psychoacoustic perceptual attribute for language as well as music processing. The perceived pitch is closely related to the periodicity of the acoustic stimulus waveform (Shofner, 2002). Therefore, pitch strength is a measure of response periodicity, indicating the robustness of neural phase locking to the stimulus F_0 contour. To determine the pitch strength of the response for the different genetic groups, autocorrelation analysis was performed. Periodicity is most often quantified by the fundamental frequency, which is the number of times the period repeats in 1 s. For quantifying the neural pitch strength, the average response of each subject was copied, the copy delayed, and the cross-correlation for the original and the delayed copy was computed by applying a short-term correlation analysis with a 40 ms sliding window and a step size of 1 ms to find the highest peak in the correlation at the time shift equal to the period (Boersma, 1993; Krishnan et al., 2005, 2010; Jeng et al., 2011a, 2011b). Correlation is maximal at time lag 0, as the signal and its copy are identical. Furthermore, in the case of a periodic signal, the correlation is high for a time shift equal to the period. Hence, periodic sounds have a peak in the correlation at time shifts corresponding to the period (and multiples of the period). In our case, a peak was found at a time lag of 10 ms, corresponding to the F_0 of the sound (frequency = 1/periodicity, so 100 Hz = 1/10 ms). To account for the transmission delay of the headphones, as well as neural delay, the analysis of the responses started at 15 ms. Pitch strength values were Fisher transformed, and averaged and analyzed with a two-way repeated-measures ANOVA with the within-subject factor condition (constant, jittered) the between-subjects factor 5-HTT expression (low and medium to high). The Greenhouse–Geisser correction was applied whenever the assumption of sphericity was violated.

Cortical responses. To obtain cortical responses, we downsampled the data to 500 Hz, reaveraged to the average reference, applied a digital bandpass filter from 0.5 to 35 Hz, and averaged the resulting long-latency auditory-evoked potentials for both timing conditions off-line in epochs ranging from –40 to 300 ms. All epochs in which the signal exceeded ± 100 Hz were rejected as containing artifacts. Seven subjects were excluded from analyses as >25% of all trials were rejected. Grand averages for both groups were computed. The resulting cortical potentials showed no prominent N1 response, probably due to the fast presentation rate (Tremblay et al., 2001), which is not optimal for the extraction of cortical

potentials due to the refractory period of neurons in the auditory cortex. We performed a two-way nonparametric permutation ANOVA with the within-subjects factor timing (constant vs jittered) and the between-subjects factor 5-HTT expression (low vs medium to high) to objectively assess whether differences between conditions and groups existed at any time point or electrode. This method prevents from incurring circular analysis of the data (double dipping) by selecting electrodes and time windows (e.g., from difference waveforms) maximizing the observed effect (Kriegeskorte et al., 2009). Also, this approach ensures that no significant differences are observed in the baseline, and it effectively corrects for multiple comparisons. We set significance to $p = 0.05$, then lowered it to $p = 0.01$ and performed 4000 permutations, applying the false discovery rate correction for multiple comparisons. There was no significant difference between groups. Further permutation analysis with a less stringent criterion (2000 and 1000 permutations instead of the initial 4000) also failed to reveal any significant effects of group on cortical potentials. As the auditory response has been shown to adapt with stimulus repetition (Haenschel et al., 2005), we performed the analyses twice, once for all trials and once for the 100 first trials. Again, as there were no main effects of the timing condition or any interaction effect, we merged the two timing conditions.

Results

We recorded the FFR and cortical responses of 79 healthy young adults to a consonant–vowel syllable (/ba/; $F_0 = 100$ Hz), presented in an either constant or jittered timing condition (Fig. 1) and in a challenging listening context, over a multispeaker background speech babble. Beforehand, we had determined the allelic variation of the 5-HTTLPR from all participants and formed two groups according to serotonin transporter expression: one group with low and one group with medium–high serotonin transporter expression. As a measure of response robustness, we analyzed the SNR of the FFR to the F_0 of the stimulus. A higher SNR implies that the neural response to the stimulus, in this case the syllable /ba/, is more precise, either due to a gain in the response to the stimulus or due to a more successful suppression of surrounding noise. Moreover, we copied part of the periodic steady-state response signal and then compared it to a time-shifted part of itself to determine periodicity, an indicator for pitch strength. To investigate whether the 5-HTTLPR polymorphism would also influence sound processing at the cortical level, we extracted and analyzed long-latency auditory-evoked potentials. A nonparametric two-way ANOVA with the factors timing (constant vs jittered) and 5-HTTLPR expression (low vs medium to high) was performed on all trials and just on the first 100 trials of each condition. Participants watched a silent movie during the experiment.

Grand average FFRs for both the constant and the jittered conditions are shown in Figure 2. We found a main effect for the between-subjects factor 5-HTT expression ($F_{(2,65)} = 5.71$; $p = 0.005$; $\eta^2 = 0.15$) with higher SNRs for the low-expression group (mean \pm SEM: low-expression group, 48.55 ± 8.02 ; medium–high-expression group, 16.63 ± 6.5) for the steady-state vowel part of the response. There was no significant difference for the consonant–vowel transition ($F_{(2,65)} = 0.64$; $p = 0.53$), nor did years of musical training influence the SNR ($F_{(1,54)} = 0.53$; $p = 0.47$). Pitch strength differed significantly between the low- and medium–high-expressing groups ($F_{(2,65)} = 4.45$; $p = 0.016$; $\eta^2 = 0.12$), with higher autocorrelation values for the low-expressing group than for the medium–high-expressing group (0.88 ± 0.06 and 0.68 ± 0.05 , respectively).

Grand average cortical potentials and scalp distribution maps for the two prominent peaks in the cortical response for both groups are shown in Figure 3. There were no significant differences between groups, with neither taking into account all trials

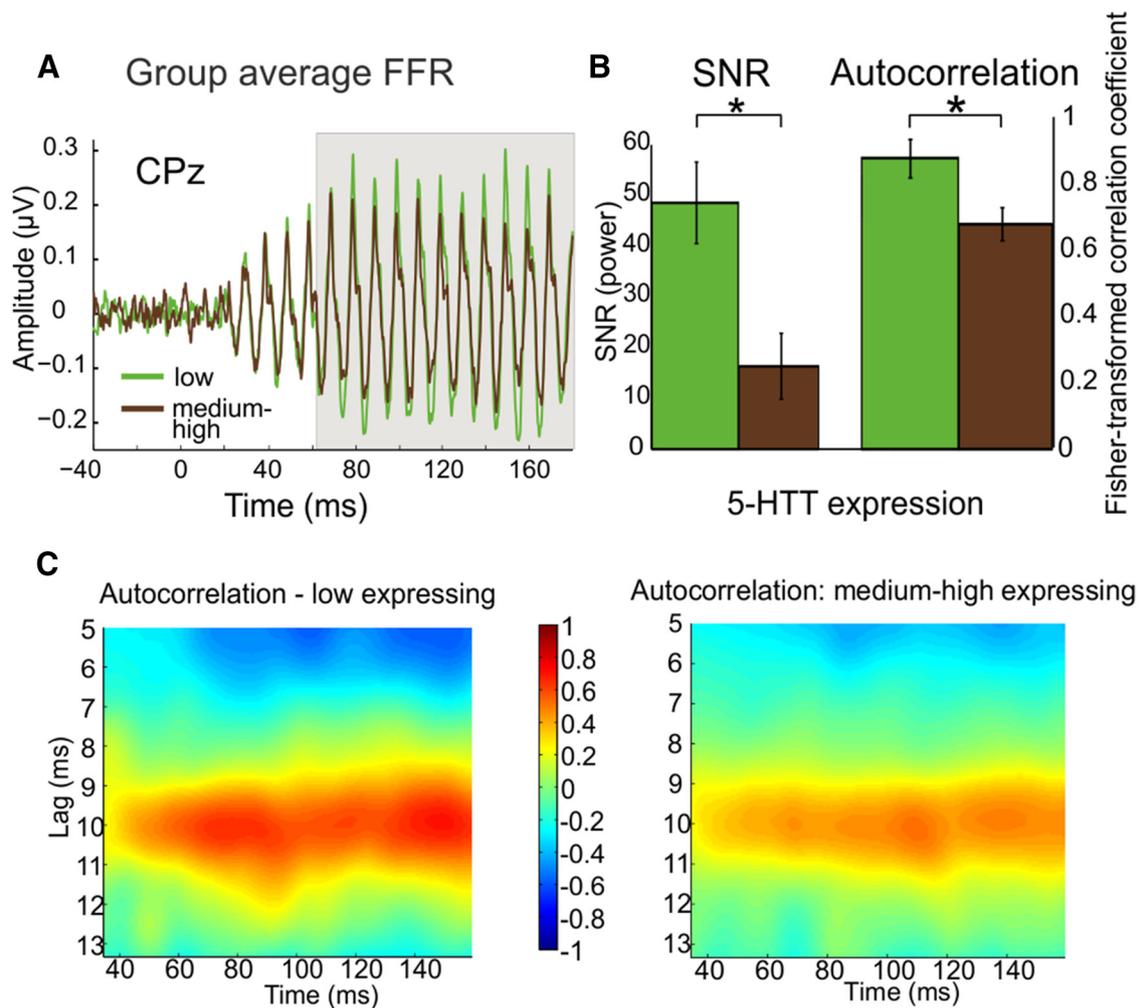


Figure 2. SNR and pitch strength results. **A**, FFR grand average for the low and medium–high 5-HTT-expressing groups. **B**, SNR and pitch strength barplots. The SNR for the low 5-HTT-expressing group was significantly higher than for the medium–high-expressing group (mean \pm SEM: low-expressing group, 48.55 ± 8.02 ; medium–high-expressing group, 16.63 ± 6.5 ; $p = 0.005$), and pitch strength was significantly higher for that group (0.88 ± 0.06 and 0.68 ± 0.05 , respectively; $p = 0.016$). **C**, Pitch strength. Autocorrelogram functions for the genetic groups. The grand average of the response for each group has been copied, and this copy was delayed to obtain autocorrelograms indicating pitch strength. The x -axis values refer to the center of each window (e.g., window 1 at 40 ms, window 2 at 41 ms), and the y -axis values refer to the time lag of the autocorrelation function. The pitch strength at the time lag of 10 ms, corresponding to the fundamental frequency of the stimulus, is significantly higher for the low-serotonin transporter expression group than for the medium–high-expression group.

nor considering only the first 100 trials. These results suggest that genetically determined 5-HTT expression influences the faithful neural representation at a subcortical level of the processing hierarchy, with independent cortical processing. However, since recording parameters, particularly the presentation rate of 3 Hz, were suboptimal for acquiring cortical potentials, caution is required in interpreting these results.

Discussion

The present study revealed that individuals with lower serotonin transporter expression, compared with individuals with medium to high expression had higher signal-to-noise ratios in the frequency following a response to the vowel of the syllable (/ba/), as well as higher pitch strength for their subcortical neuronal responses, both pointing toward a sharper speech signal extraction in the subcortical auditory pathway. The lack of significant results at cortical level indicates the specificity of the influence of 5-HTT expression to the subcortical neuronal representation of a sound, rather than its more abstract cortical processing. Although previous studies have shown the importance of serotonin for auditory perception, the 5-HTTLPR

had not been previously implicated in speech signal extraction at subcortical stages as an important mechanism for language and music perception. The influence of genetics on both speech (Morell et al., 2007) and nonspeech auditory processing (Brewer et al., 2016) has been suggested by twin studies, which have shown that auditory processing is highly determined by the individual genetic makeup.

Former studies have linked serotonin to auditory disorders, such as tinnitus (Norena et al., 1999; Simpson and Davies, 2000; Caperton and Thompson, 2011) and hyperacusis (Marriage and Barnes, 1995; Attri and Nagarkar, 2010), and the auditory symptoms of psychiatric disorders, such as an altered loudness dependence of auditory-evoked potentials in schizophrenia (Juckel et al., 2003, 2008; Park et al., 2010, 2015; Juckel, 2015), and to differential auditory brain responses in depression (Gopal et al., 2000, 2005; Chen et al., 2002; Kampf-Sherf et al., 2004; brainstem response waveforms, loudness levels, and long-latency auditory-evoked potentials N1 and P2). Variations in the serotonin level have been linked to auditory processing at the cortex, where changes in endogenous serotonin availability have been shown to

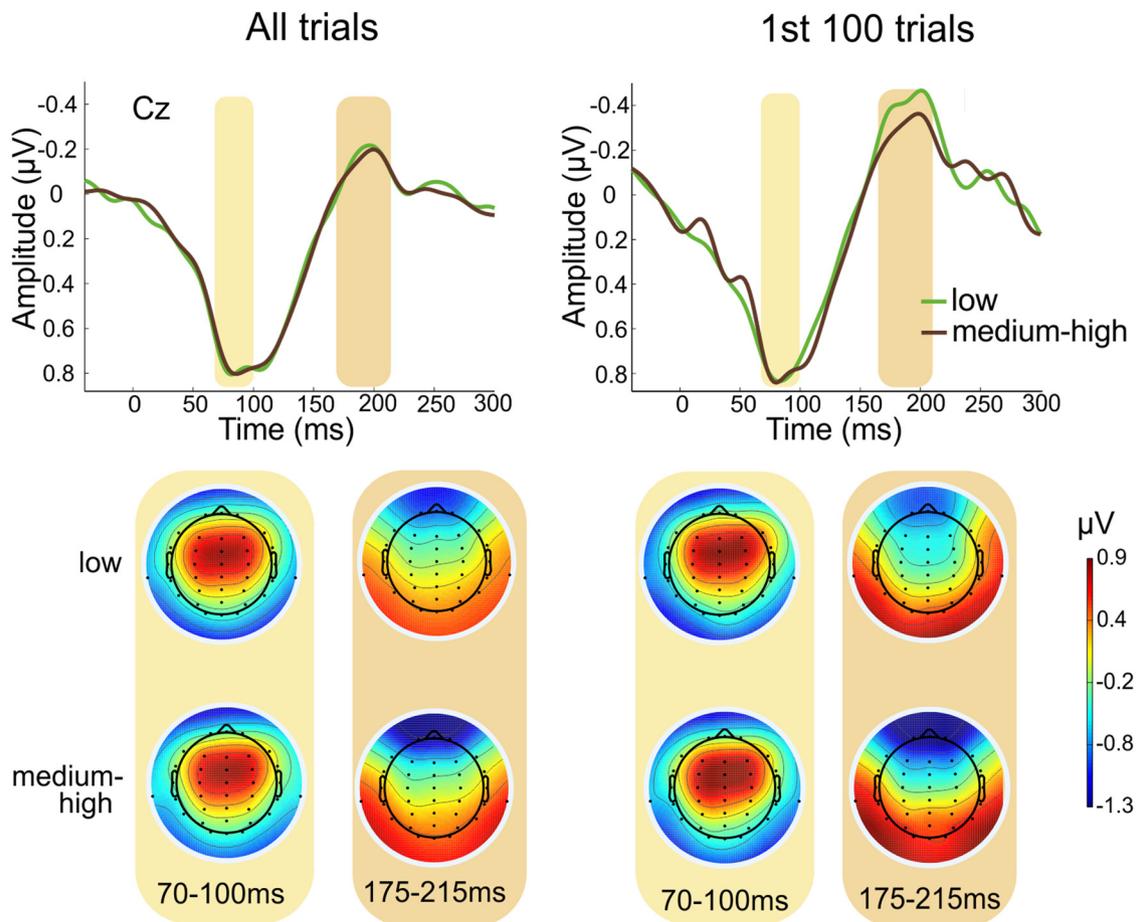


Figure 3. Long-latency cortical auditory-evoked potentials to the syllable /ba/ and scalp topographies for the two prominent peaks. On the left, results for all trials are shown, while the right part of the figure shows results for the first 100 trials only. Permutation analyses revealed no significant differences between groups at any time point or electrode.

influence the amplitude of the N1/P2 auditory-evoked component (Manjarrez et al., 2005), as well as the loudness dependence of this component (Gallinat et al., 2003; Juckel et al., 2008; Pogarell et al., 2008). Drugs related to the release of serotonin and dopamine significantly reduced P300 amplitude in a further study (Lee et al., 2016).

While we identified no previous study implying the 5-HTTLPR in speech encoding, serotonin has been linked to auditory processing downstream of the cortex, in the auditory brainstem and, particularly, in the IC. Different methodological approaches ranging from histological to immunohistochemical, to chromatographic techniques and electrical as well as chemical stimulation have corroborated the existence of serotonergic fibers in the IC of different mammalian species (Moore et al., 1978; Klepper and Herbert, 1991; Keesom and Hurley, 2016). The neuromodulator in the IC that has received the most attention is serotonin. In general, the presence of serotonin in the IC is considered essential to modulate the brain response to auditory stimuli (Hurley and Pollak, 1999, 2001, 2005; Hurley and Hall, 2011; Obara et al., 2014; Lavezzi et al., 2015). Postmortem studies in humans have shown high densities of serotonergic fibers, especially in the central nucleus of the IC (Lavezzi et al., 2015). Serotonin receptors fall into seven main families, and members of at least the 5-HT1 (Chalmers and Watson, 1991; Pompeiano et al., 1992; Thompson et al., 1994; Wright et al., 1995), 5-HT2 (Wright et al., 1995), 5-HT3 (Bohorquez and Hurley, 2009), 5-HT4 (Waeber et al., 1994), and 5-HT7 (To et al., 1995) families have been found in auditory neurons of the IC. There is suggestive evidence that serotonin influences acoustic processing in the IC by modulating the

tone of activity of GABAergic inhibitory interneurons (Peruzzi and Dut, 2004; Obara et al., 2014). When serotonin is applied endogenously or exogenously, usually a suppression or decrease in the amplitude of evoked responses and spontaneous activity in the IC is observed (Hurley and Pollak, 1999). Functionally, this suppression might result in an increased selectivity of responses to auditory stimuli (Hurley and Sullivan, 2012), possibly leading to a more robust neural response by tuning synaptic activity to the stimulus properties and decreasing the representation of noise, as observed in the present study.

Coffey et al. (2016) suggested in a recent article that although the FFR is originated mainly in subcortical nuclei, there might exist cortical contributions to the signal as well. The F_0 of 100 Hz of the syllable used in this study might well be within the phase-locking capabilities of single neurons of the cortex. Nevertheless, although cortical activity might contribute to the FFR, we found no difference between groups in the cortical response to the stimulus, indicating that genetically mediated 5-HTT expression possibly only affects the auditory processing at an early, subcortical level. This should be confirmed, however, by future studies with a slower stimulus presentation rate, which is more suitable for the extraction of cortical potentials.

It has been stated that individual differences in auditory function can be observed even in typically developing individuals (Chandrasekaran et al., 2012; Hairston et al., 2013; Skoe et al., 2013b), and the ultimate determinants for subcortical differences may arise from a combination of environmental and genetic factors (Chandrasekaran et al., 2012). Different models of how

acoustic experience and auditory training mold the subcortical processing of sound have been proposed (Krishnan and Gandour, 2009; Kraus and Chandrasekaran, 2010; Patel, 2011). These converge in the so-called “layering hypothesis” (Skoe and Chandrasekaran, 2014), a recent proposal for how auditory function is shaped by different types of experience at different time scales (from seconds to years) and taking into account the metaplasticity of the system. The layering hypothesis states that not only cortical, but also subcortical auditory processing is influenced by the specific auditory experience of each individual, resulting in individual differences in auditory function (Chandrasekaran et al., 2012; Hairston et al., 2013; Skoe et al., 2013b). Our results show that common variants in the 5-HTTLPR influence the robustness and accuracy of speech encoding at subcortical stages, and could therefore constrain or expand the extent to which experience can shape the structure and function of the subcortical auditory pathway. The scope of this influence yet has to be determined, and eventually may entail a complex interplay both with other genes and the environmental factors that have been disclosed so far (Skoe and Chandrasekaran, 2014). Because a large range of studies has shown that encoding accuracy at subcortical stages is experience dependent and can be trained (for review, see Skoe and Chandrasekaran, 2014), special attention might be given at an early developmental stage to individuals with medium to high 5-HTT expression, a possible genetically determined difficulty for auditory processing that may be counterbalanced by musical or other specific auditory training. This may be of interest, for example, regarding the auditory processing disorder, in which the major symptom is a difficulty in efficiently extracting speech from noise despite normal hearing sensitivity (Rosen et al., 2010). The prevalence of this disorder is especially high in the pediatric population (in combination with co-occurring learning disorders, it is estimated to be as high as 10%; Brewer et al., 2016), and the identification of the genetic and neurophysiologic underpinnings might significantly contribute to the understanding of the disorder and yield to early intervention. Our study identifies the 5-HTTLPR as being implicated in auditory encoding. Whether it might be involved in the genetic predisposition to the disorder, eventually in combination with other genes and influenced by acoustic experience, cannot be answered by the present study, and remains an interesting question for future studies.

On the other hand, it is of note that a number of studies have revealed cognitive advantages for individuals carrying the short allele of the 5-HTTLPR (Roiser et al., 2007; Borg et al., 2009; Enge et al., 2011), which in the present study were shown to have a more accurate speech sound encoding at subcortical stages. Since a relation between the auditory brainstem responses, including the FFR, and higher-order cognitive functions, such as rapid learning (Skoe et al., 2013b), reading abilities (Hornickel and Kraus, 2013), and executive function (Krizman et al., 2012), has been established, we are tempted to suggest the FFR to complex speech sounds as an intermediate phenotype not only for auditory processing, but eventually for a range of specific cognitive assets.

We believe our results pave the way for a series of questions for future research. As it has been shown that the FFR is plastic and experience dependent, it would be interesting to study the potential interactions between carrying specific allelic variants of the 5-HTTLPR and musical experience or bilingualism, for instance. In addition, it may turn out that the 5-HTTLPR is implicated in specific clinical groups involving several types of speech impairments, such as in stuttering, which has been linked not only to

production, but also to auditory perception (Corbera et al., 2005) and dyslexia. While awaiting future confirmation, our results point to a higher ratio of individuals with medium to high serotonin transporter expression in these groups than in the general population. Likewise, low 5-HTT expression might result in an advantage for musical ability or a better achievement in the correct pronunciation of a newly learned foreign language.

To conclude, we found a significantly higher SNR and pitch strength of the subcortical auditory response to speech sounds in individuals with a low serotonin transporter expression, compared with those with medium to high expression, with both measures pinpointing toward more accurate speech encoding at subcortical stages of the auditory pathway in the former group. While experience-dependent plasticity of subcortical sound and speech encoding has been shown in an established line of research, the involvement of the 5-HTTLPR polymorphism adds a new layer to the understanding of how the interplay of different hardwired and experiential factors shapes a function as complex as speech encoding. Future studies should aim at establishing the extent to which auditory processing at subcortical stages is influenced by the genetic makeup, and how genetic expression can be modulated by experience-dependent plasticity.

References

- Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245:35–47. [CrossRef Medline](#)
- Assmann PF, Summerfield AQ (2004) The perception of speech under adverse conditions. In: *Speech processing in the auditory system* (Greenberg S, Ainsworth WA, Popper AN, Fay RR, eds), pp 231–308. New York: Springer.
- Attri D, Nagarkar AN (2010) Resolution of hyperacusis associated with depression, following lithium administration and directive counselling. *J Laryngol Otol* 124:919–921. [CrossRef Medline](#)
- Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Reading and subcortical auditory function. *Cereb Cortex* 19:2699–2707. [CrossRef Medline](#)
- Basu M, Krishnan A, Weber-Fox C (2010) Brainstem correlates of temporal auditory processing in children with specific language impairment. *Dev Sci* 13:77–91. [CrossRef Medline](#)
- Boersma J (1993) Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc Inst Phonetic Sci* 17:97–110.
- Bohorquez A, Hurley LM (2009) Activation of serotonin 3 receptors changes in vivo auditory responses in the mouse inferior colliculus. *Hear Res* 251:29–38. [CrossRef Medline](#)
- Borg J, Henningson S, Saijo T, Inoue M, Bah J, Westberg L, Lundberg J, Jovanovic H, Andrée B, Nordstrom AL, Halldin C, Eriksson E, Farde L (2009) Serotonin transporter genotype is associated with cognitive performance but not regional 5-HT 1 A receptor binding in humans. *Int J Neuropsychopharmacol* 12:783–792. [CrossRef Medline](#)
- Brewer CC, Zalewski CK, King KA, Zobay O, Riley A, Ferguson MA, Bird JE, McCabe MM, Hood LJ, Drayna D, Griffith AJ, Morell RJ, Friedman TB, Moore DR (2016) Heritability of non-speech auditory processing skills. *Eur J Hum Genet* 24:1137–1144. [CrossRef Medline](#)
- Caperton KK, Thompson AM (2011) Activation of serotonergic neurons during salicylate-induced tinnitus. *Otol Neurotol* 32:301–307. [CrossRef Medline](#)
- Chalmers DT, Watson SJ (1991) Comparative anatomical distribution of 5-HT1A receptor mRNA and 5-HT1A binding in rat brain—a combined in situ hybridisation/in vitro receptor autoradiographic study. *Brain Res* 561:51–60. [CrossRef Medline](#)
- Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47:236–246. [CrossRef Medline](#)
- Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N (2009) Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* 64:311–319. [CrossRef Medline](#)
- Chandrasekaran B, Kraus N, Wong PC (2012) Human inferior colliculus

- activity relates to individual differences in spoken language learning. *J Neurophysiol* 107:1325–1336. [CrossRef Medline](#)
- Chen TJ, Yu YW, Chen MC, Tsai SJ, Hong CJ (2002) Association analysis for serotonin transporter promoter polymorphism and auditory evoked potentials for major depression. *Neuropsychobiology* 46:57–60. [CrossRef Medline](#)
- Coffey EBJ, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7:11070. [CrossRef Medline](#)
- Cooley JW, Tukey JW (1964) An algorithm for the machine calculation complex Fourier series. *Int J Comput Math* 19:297–301.
- Corbera S, Corral MJ, Escera C, Idiazábal MA (2005) Abnormal speech sound representation in persistent developmental stuttering. *Neurology* 65:1246–1252. [CrossRef Medline](#)
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21. [CrossRef Medline](#)
- Delorme A, Sejnowski T, Makeig S (2007) Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34:1443–1449. [CrossRef Medline](#)
- Díaz B, Baus C, Escera C, Costa A, Sebastián-Gallés N (2008) Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc Natl Acad Sci U S A* 105:16083–16088. [CrossRef Medline](#)
- Enge S, Fleischhauer M, Lesch KP, Strobel A (2011) On the role of serotonin and effort in voluntary attention: evidence of genetic variation in N1 modulation. *Behav Brain Res* 216:122–128. [CrossRef Medline](#)
- Gallinat J, Senkowski D, Wernicke C, Juckel G, Becker J, Sander T, Smolka M, Hegerl U, Rommelspacher H, Winterer G, Herrmann WM (2003) Allelic variants of the functional promoter polymorphism of the human serotonin transporter gene is associated with auditory cortical stimulus processing. *Neuropsychopharmacology* 28:530–532. [CrossRef Medline](#)
- Gopal KV, Daly DM, Daniloff RG, Pennartz L (2000) Effects of selective serotonin reuptake inhibitors on auditory processing: case study. *J Am Acad Audiol* 11:454–463. [Medline](#)
- Gopal KV, Briley KA, Goodale ES, Hendea OM (2005) Selective serotonin reuptake inhibitors treatment effects on auditory measures in depressed female subjects. *Eur J Pharmacol* 520:59–69. [CrossRef Medline](#)
- Haenschel C, Vernon DJ, Dwivedi P, Gruzeliel JH, Baldeweg T (2005) Event-related brain potential correlates of human auditory sensory memory-trace formation. *J Neurosci* 25:10494–10501. [CrossRef Medline](#)
- Hairston WD, Letowski TR, McDowell K (2013) Task-related suppression of the brainstem frequency following response. *PLoS One* 8:e55215. [CrossRef Medline](#)
- Heils A, Teufel A, Petri S, Stöber G, Riederer P, Bengel D, Lesch KP (1996) Allelic variation of human serotonin transporter gene expression. *J Neurochem* 66:2621–2624. [Medline](#)
- Hornickel J, Kraus N (2013) Unstable representation of sound: a biological marker of dyslexia. *J Neurosci* 33:3500–3504. [CrossRef Medline](#)
- Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci U S A* 106:13022–13027. [CrossRef Medline](#)
- Hornickel J, Chandrasekaran B, Zecker S, Kraus N (2011) Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. *Behav Brain Res* 216:597–605. [CrossRef Medline](#)
- Hu X, Oroszi G, Chun J, Smith TL, Goldman D, Schuckit MA (2005) An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. *Alcohol Clin Exp Res* 29:8–16. [CrossRef Medline](#)
- Hurley LM, Hall IC (2011) Context-dependent modulation of auditory processing by serotonin. *Hear Res* 279:74–84. [CrossRef Medline](#)
- Hurley LM, Pollak GD (1999) Serotonin differentially modulates responses to tones and frequency-modulated sweeps in the inferior colliculus. *J Neurosci* 19:8071–8082. [Medline](#)
- Hurley LM, Pollak GD (2001) Serotonin effects on frequency tuning of inferior colliculus neurons. *J Neurophysiol* 85:828–842. [Medline](#)
- Hurley LM, Pollak GD (2005) Serotonin shifts first-spike latencies of inferior colliculus neurons. *J Neurosci* 25:7876–7886. [CrossRef Medline](#)
- Hurley LM, Sullivan MR (2012) From behavioral context to receptors: serotonergic modulatory pathways in the IC. *Front Neural Circuits* 6:58. [CrossRef Medline](#)
- Jafari Z, Malayeri S, Rostami R (2015) Subcortical encoding of speech cues in children with attention deficit hyperactivity disorder. *Clin Neurophysiol* 126:325–332. [CrossRef Medline](#)
- Jeng FC, Chung HK, Lin CD, Dickman B, Hu J (2011a) Exponential modeling of human frequency-following responses to voice pitch. *Int J Audiol* 50:582–593. [CrossRef Medline](#)
- Jeng FC, Hu J, Dickman B, Lin CY, Lin CD, Wang CY, Chung HK, Li X (2011b) Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. *Int J Audiol* 50:14–26. [CrossRef Medline](#)
- Juckel G (2015) Serotonin: from sensory processing to schizophrenia using an electrophysiological method. *Behav Brain Res* 277:121–124. [CrossRef Medline](#)
- Juckel G, Gallinat J, Riedel M, Sokullu S, Schulz C, Möller HJ, Müller N, Hegerl U (2003) Serotonergic dysfunction in schizophrenia assessed by the loudness dependence measure of primary auditory cortex evoked activity. *Schizophr Res* 64:115–124. [CrossRef Medline](#)
- Juckel G, Gudlowski Y, Müller D, Özgürdal S, Brüne M, Gallinat J, Frodl T, Witthaus H, Uhl I, Wutzler A, Pogarell O, Mülert C, Hegerl U, Meisenzahl EM (2008) Loudness dependence of the auditory evoked N1/P2 component as an indicator of serotonergic dysfunction in patients with schizophrenia—a replication study. *Psychiatry Res* 158:79–82. [CrossRef Medline](#)
- Kampf-Sherf O, Zlotogorski Z, Gilboa A, Speedie L, Lereya J, Rosca P, Shavit Y (2004) Neuropsychological functioning in major depression and responsiveness to selective serotonin reuptake inhibitors antidepressants. *J Affect Disord* 82:453–459. [CrossRef Medline](#)
- Keesom SM, Hurley LM (2016) Socially induced serotonergic fluctuations in the male auditory midbrain correlate with female behavior during courtship. *J Neurophysiol* 115:1786–1796. [CrossRef Medline](#)
- Klatt DH (1980) Software for cascade/parallel formant synthesizer. *J Acoust Soc Am* 67:971–975. [CrossRef](#)
- Klepper A, Herbert H (1991) Distribution and origin of noradrenergic and serotonergic fibers in the cochlear nucleus and inferior colliculus of the rat. *Brain Res* 557:190–201. [CrossRef Medline](#)
- Kraus N, Chandrasekaran B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605. [CrossRef Medline](#)
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540. [CrossRef Medline](#)
- Krishnan A, Gandour JT (2009) The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain Lang* 110:135–148. [CrossRef Medline](#)
- Krishnan A, Xu Y, Gandour J, Cariani P (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 25:161–168. [CrossRef Medline](#)
- Krishnan A, Bidelman GM, Gandour JT (2010) Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hear Res* 268:60–66. [CrossRef Medline](#)
- Krizman J, Marian V, Shook A, Skoe E, Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci U S A* 109:7877–7881. [CrossRef Medline](#)
- Lavezzi AM, Pusioli T, Matturri L (2015) Cytoarchitectural and functional abnormalities of the inferior colliculus in sudden unexplained perinatal death. *Medicine (Baltimore)* 94:e487. [CrossRef Medline](#)
- Lee H, Wang GY, Curley LE, Sollers JJ, Kydd RR, Kirk IJ, Russell BR (2016) Acute effects of BZP, TFMPP and the combination of BZP and TFMPP in comparison to dexamphetamine on an auditory oddball task using electroencephalography: a single-dose study. *Psychopharmacology (Berl)* 233:863–871. [CrossRef Medline](#)
- Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Müller CR, Hamer DH, Murphy DL (1996) Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 274:1527–1531. [CrossRef Medline](#)
- Manjarrez G, Hernandez E, Robles A, Hernandez J (2005) N1/P2 component of auditory evoked potential reflect changes of the brain serotonin biosynthesis in rats. *Nutr Neurosci* 8:213–218. [CrossRef Medline](#)
- Marriage J, Barnes NM (1995) Is central hyperacusis a symptom of 5-hydroxytryptamine (5-HT) dysfunction? *J Laryngol Otol* 109:915–921. [Medline](#)
- Møller AR, Jannetta PJ, Sekhar LN (1988) Contributions from the auditory nerve to the brain-stem auditory evoked potentials (BAEPs): results of intracranial recording in man. *Electroencephalogr Clin Neurophysiol* 71:198–211. [CrossRef Medline](#)

- Moore RY, Halaris AE, Jones BE (1978) Serotonin neurons of the midbrain raphe: ascending projections. *J Comp Neurol* 180:417–438. [CrossRef Medline](#)
- Morell RJ, Brewer CC, Ge D, Snieder H, Zalewski CK, King KA, Drayna D, Friedman TB (2007) A twin study of auditory processing indicates that dichotic listening ability is a strongly heritable trait. *Hum Genet* 122:103–111. [CrossRef Medline](#)
- Norena A, Cransac H, Chéry-Croze S (1999) Towards an objectification by classification of tinnitus. *Clin Neurophysiol* 110:666–675. [CrossRef Medline](#)
- Obara N, Kamiya H, Fukuda S (2014) Serotonergic modulation of inhibitory synaptic transmission in mouse inferior colliculus. *Biomed Res* 35: 81–84. [CrossRef Medline](#)
- Pallier C, Bosch L, Sebastián-Gallés N (1997) A limit on behavioral plasticity in speech perception. *Cognition* 64:B9–B17. [CrossRef Medline](#)
- Parbery-Clark A, Strait DL, Hittner E, Kraus N (2013) Musical training enhances neural processing of binaural sounds. *J Neurosci* 33:16741–16747. [CrossRef Medline](#)
- Park YM, Lee SH, Kim S, Bae SM (2010) The loudness dependence of the auditory evoked potential (LDAEP) in schizophrenia, bipolar disorder, major depressive disorder, anxiety disorder, and healthy controls. *Prog Neuropsychopharmacol Biol Psychiatry* 34:313–316. [CrossRef Medline](#)
- Park YM, Jung E, Kim HS, Hahn SW, Lee SH (2015) Differences in central serotonergic transmission among patients with recent onset, subchronic, and chronic schizophrenia as assessed by the loudness dependence of auditory evoked potentials. *Schizophr Res* 168:180–184. [CrossRef Medline](#)
- Patel AD (2011) Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front Psychol* 2:142. [CrossRef Medline](#)
- Peruzzi D, Dut A (2004) GABA, serotonin and serotonin receptors in the rat inferior colliculus. *Brain Res* 998:247–250. [CrossRef Medline](#)
- Pogarell O, Koch W, Schaaff N, Pöpperl G, Mulert C, Juckel G, Möller HJ, Hegerl U, Tsuchi K (2008) [123I] ADAM brainstem binding correlates with the loudness dependence of auditory evoked potentials. *Eur Arch Psychiatry Clin Neurosci* 258:40–47. [CrossRef Medline](#)
- Pompeiano M, Palacios JM, Mengod G (1992) Distribution and cellular localization of mRNA coding for 5-HT_{1A} receptor in the rat brain: correlation with receptor binding. *J Neurosci* 12:440–453. [Medline](#)
- Rocha-Muniz CN, Befi-Lopes DM, Schochat E (2012) Investigation of auditory processing disorder and language impairment using the speech-evoked auditory brainstem response. *Hear Res* 294:143–152. [CrossRef Medline](#)
- Roiser JP, Müller U, Clark L, Sahakian BJ (2007) The effects of acute tryptophan depletion and serotonin transporter polymorphism on emotional processing in memory and attention. *Int J Neuropsychopharmacol* 10: 449–461. [CrossRef Medline](#)
- Rosen S, Cohen M, Vanniasegaram I (2010) Auditory and cognitive abilities of children suspected of auditory processing disorder (APD). *Int J Pediatr Otorhinolaryngol* 74:594–600. [CrossRef Medline](#)
- Russo NM, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A, Kraus N (2008) Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clin Neurophysiol* 119:1720–1731. [CrossRef Medline](#)
- Shofner WP (2002) Perception of the periodicity strength of complex sounds by the chinchilla. *Hear Res* 173:69–81. [CrossRef Medline](#)
- Simpson JJ, Davies WE (2000) A review of evidence in support of a role for 5-HT in the perception of tinnitus. *Hear Res* 145:1–7. [CrossRef Medline](#)
- Skoe E, Chandrasekaran B (2014) The layering of auditory experiences in driving experience-dependent subcortical plasticity. *Hear Res* 311:36–48. [CrossRef Medline](#)
- Skoe E, Kraus N (2010a) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31:302–324. [CrossRef Medline](#)
- Skoe E, Kraus N (2010b) Hearing it again and again: on-line subcortical plasticity in humans. *PLoS One* 5:e13645. [CrossRef Medline](#)
- Skoe E, Krizman J, Kraus N (2013a) The impoverished brain: disparities in maternal education affect the neural response to sound. *J Neurosci* 33: 17221–17231. [CrossRef Medline](#)
- Skoe E, Krizman J, Spitzer E, Kraus N (2013b) The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience* 243:104–114. [CrossRef Medline](#)
- Slabu L, Grimm S, Escera C (2012) Novelty detection in the human auditory brainstem. *J Neurosci* 32:1447–1452. [CrossRef](#)
- Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol* 39:465–472. [CrossRef Medline](#)
- Song JH, Skoe E, Banai K, Kraus N (2011) Perception of speech in noise: neural correlates. *J Cogn Neurosci* 23:2268–2279. [CrossRef Medline](#)
- Spielberger CD, Gorsuch RL, Lushene RE (1988) *Cuestionario de ansiedad estado-rasgo—manual, Ed 3*. Madrid, Spain: TEA Ediciones (by permission of Consulting Psychologists Press, Sunnyvale, CA).
- Thompson GC, Thompson AM, Garrett KM, Britton BH (1994) Serotonin and serotonin receptors in the central auditory system. *Otolaryngol Head Neck Surg* 110:93–102. [Medline](#)
- To ZP, Bonhaus DW, Eglén RM, Jakeman LB (1995) Characterization and distribution of putative 5-HT₇ receptors in guinea-pig brain. *Br J Pharmacol* 115:107–116. [CrossRef Medline](#)
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B (2001) Central auditory plasticity: changes in the N1–P2 complex after speech-sound training. *Ear Hear* 22:79–90. [CrossRef Medline](#)
- Waeber C, Sebben M, Nieoullon A, Bockaert J, Dumuis A (1994) Regional distribution and ontogeny of 5-HT₄ binding sites in rodent brain. *Neuropharmacology* 33:527–541. [CrossRef Medline](#)
- White-Schwoch T, Woodruff Carr K, Thompson EC, Anderson S, Nicol T, Bradlow AR, Zecker SG, Kraus N (2015) Auditory processing in noise: a preschool biomarker for literacy. *PLoS Biol* 13:e1002196. [CrossRef Medline](#)
- Wright DE, Seroogy KB, Lundgren KH, Davis BM, Jennes L (1995) Comparative localization of serotonin_{1A}, _{1C}, and ₂ receptor subtype mRNAs in rat brain. *J Comp Neurol* 351:357–373. [CrossRef Medline](#)