FREQUENCY FOLLOWING RESPONSES IN ADULTS WHO CAN OR CANNOT SING IN TUNE

Milaine Dominici Sanfins[1,2,A-G], **Ingrid Gielow**[3,A,D-E], **Glaucya Madazio**[3,A,D], **Francine Honorio**[3,A], **Tatiana Bordin**[4,C-D], **Magdalena B. Skarzynska**[5,E], **Mara Behlau**[6,7,A,E-F]

[1] Electrophysiology, Centro de Eletrofisiologia e Neuroaudiologia Avançada, Brazil
[2] Hearing, Centro de Estudos da Voz (CEV), Brazil
[3] Hearing, Centro de Estudos da Voz (CEV), São Paulo, Brazil, Brazil
[4] Estatistical, Instituto de pesquisa Eldorado, Brazil
[5] Institute of Physiology and Pathology of Hearing, World Hearing Center, Poland
[6] Voice, Centro de Estudos da Voz (CEV), Brazil
[7] Voice, Universidade Federal de São Paulo, Brazil

Corresponding author: Milaine Dominici Sanfins; Electrophysiology, Centro de Eletrofisiologia e Neuroaudiologia Avançada, Avenida Jacutinga, 220- apto 12, 04515-030, São Paulo, Brazil; email: msanfins@uol.com.br; Phone: +5511990033092

Contributions:
A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

Abstract

Background: It is believed that auditory processing occurs normally in people who can sing in tune and improperly in people who cannot. Auditory feedback seems to be a crucial factor in the way the voice is produced and monitored. Evaluation of auditory processing using the Frequency Following Response (FFR) allows fine-grained neural processing to be objectively identified and might be a way of differentiating between those who sing in-tune and those who sing off-tune. The aim of this study was to analyse the FFR responses of people who can sing in-tune and compare them to those who sing off-tune.

Material and methods: FFR responses were recorded in 37 adults who were assigned to one of two groups: (i) a control group (CG) consisting of 17 adults who could sing in-tune (ii) an experimental group (EG) consisting of 20 adults who sang off-tune.

Results: There were statistically significant differences in the electro-physiological responses of the EG compared to the CG for the latencies of waves A, C, D, and F in the right ear. In contrast, FFR amplitude measurements did not seem to be a suitable parameter for identifying changes in the coding of speech sounds.

Conclusions: FFR responses in the EG showed a different pattern from the CG group, with a number of longer latencies in the EG. However, FFR amplitude did not differ significantly between the groups.

Key words: adults • hearing • voice • speech perception • auditory event-related potentials (ERPs)

CZĘSTOTLIWOŚCIOWE ODPOWIEDZI WYWOŁANE U DOROSŁYCH, KTÓRZY MOGĄ LUB NIE MOGĄ ŚPIEWAĆ CZYSTO

Streszczenie

Wprowadzenie: Uważa się, że u osób, które potrafią czysto śpiewać, przetwarzanie słuchowe jest prawidłowe, a u osób, które nie mają takiej umiejętności - nieprawidłowe. Odpowiedzi słuchowe wydają się kluczowym czynnikiem w sposobie wytwarzania i monitorowania głosu. Ocena przetwarzania słuchowego za pomocą częstotliwościowych odpowiedzi wywołanych (FFR) pozwala na obiektywną identyfikację drobnoziarnistego przetwarzania nerwowego i może być sposobem na rozróżnienie między tymi, którzy śpiewają czysto, a tymi, którzy nie śpiewają czysto. Celem tego badania była analiza odpowiedzi FFR u osób, którzy potrafią śpiewać czysto i porównanie ich z tymi, którzy takiej umiejętności nie posiadają.

Materiał i metody: Odpowiedzi FFR zbadano u 37 osób dorosłych, którzy zostali przydzieleni do jednej z dwóch grup: (i) grupa kontrolna (CG) składająca się z 17 osób dorosłych, którzy potrafili śpiewać czysto (ii) grupa eksperymentalna (EG) składająca się z 20 osób dorosłych nie śpiewających czysto.

Wyniki: Wystąpiły istotne statystycznie różnice w odpowiedziach elektrofizjologicznych w grupie EG w porównaniu do grupy CG dla latencji fal A, C, D i F w prawym uchu. Natomiast amplitudy FFR nie okazały się odpowiednim parametrem do identyfikacji zmian w kodowaniu dźwięków mowy.

Wnioski: Odpowiedzi FFR w grupie EG wykazywały inny wzór niż w grupie CG – w grupie EG zaobserwowano pewną liczbę dłuższych latencji. Jednak amplituda FFR nie różniła się istotnie między grupami.

Słowa kluczowe: dorośli • słuch • głos • percepcja mowy • słuchowe potencjały wywołane (ERP)

Introduction

Music and language are human abilities but their processing by the brain differs. Music is processed predominantly by the right hemisphere, mainly concerning prosodic and melodic features [1], while language is processed by specific brain structures largely located in the left hemisphere. So although music and language are processed by distinct systems having different features, they operate with similar auditory perception mechanisms [2].

In terms of voice, monitoring of vocal quality production is carried out by the central nervous system through hearing. Thus, the voice is directly related to auditory processing. Malfunctioning in this area might cause difficulties in controlling the frequency and/or the intensity of the voice. If such a loss of control occurs, then if a person intends to sing a certain note the result is off-pitch; here we call it off-tune singing [3]. In severe cases, the condition is called amusia or tone-deafness [3]. The deviation from correct pitch is a complex process and may occur from a lack of exposure to music, which is one obvious cause. However, it is presently not possible to pin-point where the reference pitch generator is located. There are two aspects related to this problem: difficulties in sound perception and/or vocal production [4], although other authors point to other sources such as memory and language [5]. In-tune individuals seem to have correct processing and perception of sounds, while off-tune individuals have impairment in these skills [6[.

Experience and training in music, either short or long term, causes changes in how the central auditory nervous system (CANS) operates. It is well known that musicians, compared to non-musicians, have enhanced sensitivity to their instrument and more robust representation of musical notes, both in the auditory brainstem and auditory cortex. In addition, musicians have better auditory processing skills [7-10], motor learning, pitch and melodic perception, vocabulary, verbal and nonverbal memory [11-13], finer perceptual qualities, and better neural encoding of speech stimuli [8, 11]. In addition, recent studies report a correlation between musical learning and language development. Thus, it is believed that an individual with a good knowledge of music and extensive musical experience is more likely to have a more well-developed set of language skills [14].

The literature suggests that dysfunction of auditory processing or neurophysiology can be identified objectively based on responses obtained from auditory evoked potentials (AEPs) [15, 16]. The Frequency Following Response (FFR) appears to be an instrument that may be able to gauge how well the neural coding process is functioning. The FFR seems to provide information about the temporal and spectral encoding of sounds in the auditory system through three main components: frequency, structure of formants, and timing. In addition, the FFR allows one to analyse time and prosody cues, which are important for discriminating consonants and vowels [17]. It also gives information on frequency cues, which are important for understanding intonation [18, 19]. Any difficulty in processing some of these aspects will lead to difficulties in auditory processing and the perception of speech sounds [20-23].

The FFR is an electrophysiological test that can assess the integrity with which verbal sounds are processed in subcortical and cortical regions. It is a fast, objective method and is considered a good measure of communication functioning and can be applied in different populations [7, 9, 15]. However, there has been no study that measures the difference in FFR responses between individuals who can sing in-tune and those who sing off-tune. The aim of this study was to investigate the effect of this factor on FFR responses.

Materials and methods

Statement of ethics

This study was approved by the Committee for Ethics in Research under protocol number 1.191.303 at the CAEE: 41305515.9.0000.5511. Informed consent for research was obtained in writing from all participants after an explanation of the nature, purpose, and expected results of the study.

Participants

A total of 37 individuals participated in this study, 20 female and 17 male, aged between 20 and 57 years. The subjects were divided into two groups:

(i) A control group (CG) consisted of 17 professional musicians (10 females and 7 males) who could sing in-tune.
(ii) The experimental group (EG) consisted of 20 non-musicians (10 females and 10 males) who sang off-tune.

Inclusion criteria

The inclusion criteria were defined as:

(i) Control Group (CG)

- Air-conduction threshold below 20 dB HL for octaves from 250 to 8000 Hz;
- Bone-conduction thresholds below 15 dB HL for octaves between 500 to 4000 Hz;
- Type A tympanogram with compliance between 0.3 and 1.3 mmhos and pressure between −100 and +200 daPa and associated with the presence of ipsilateral and contralateral acoustic reflexes in both ears;
- A click Auditory Brainstem Response (ABR) with waves I, III, and V present and with an inter-peak interval I–III, III–V, and I–V within normal standards in both ears;
- No syndromic hearing impairment, or other middle ear or inner ear diseases;
- No current or prior neurological, cognitive, or psychiatric disorders;
- No complaint of learning or speech disorder;
- Professional musician without tuning anomalies as confirmed by administration of a pitch-matching test (see below).

(ii) Experimental group (EG)

- Air-conduction threshold below 20 dB HL for octaves from 250 to 8000 Hz;

- Bone-conduction thresholds below 15 dB HL for octaves between 500 to 4000 Hz;
- Type A tympanogram with compliance between 0.3 and 1.3 mmhos and pressure between −100 and +200 daPa associated with the presence of ipsilateral and contralateral acoustic reflexes in both ears;
- A click Auditory Brainstem Response (ABR) with waves I, III, and V present and with an inter-peak interval I– III, III–V, and I–V within normal standards in both ears;
- No syndromic hearing impairment, or other middle ear or inner ear disease;
- No current or prior neurological, cognitive, or psychiatric disorder;
- No complaint of learning or speech disorder;
- No musical ability, with errors in tuning ability confirmed by a pitch-matching test.

Procedures

Audiological evaluation

a. Audiometric evaluation was performed via air conduction at 0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz and bone conduction at 0.5, 1, 2, and 4 kHz. Normal auditory threshold was considered to be normal up to 15 dB for bone conduction and up to 20 dB for air conduction according to the classification of Davis and Silverman [24]. Testing was performed using an Interacoustics AC 40 audiometer.

b. Speech Recognition Threshold (SRT). A list of disyllables was adopted and the final result was the intensity at which the participant scored 50% of the words presented.

c. Speech Recognition Index (SRI) was tested at 40 dB above the mean tonal threshold of 0.5, 1, and 2 kHz using a list of monosyllabic words. SRI was considered normal if the percentage of correct answers was between 88 and 100%.

d. Immittanciometry (tympanometry and acoustic reflex). Tympanometry was performed with a 226 Hz probe tone. Ipsilateral and contralateral acoustic reflexes were probed at frequencies of 0.5, 1, 2, and 4 kHz. Normal subjects presented a peak maximum compliance at atmospheric pressure (0 daPa) and an equivalent volume of 0.3 to 1.3 mL according to the proposal of Jerger (1970) [25]. Immittanciometry was performed using the Interacoustics Immitator 235h. All equipment was calibrated according to ISO-389 and IEC-645 standards. Subjects who had normal responses in the basic audiological evaluation were then tested by auditory electrophysiology.
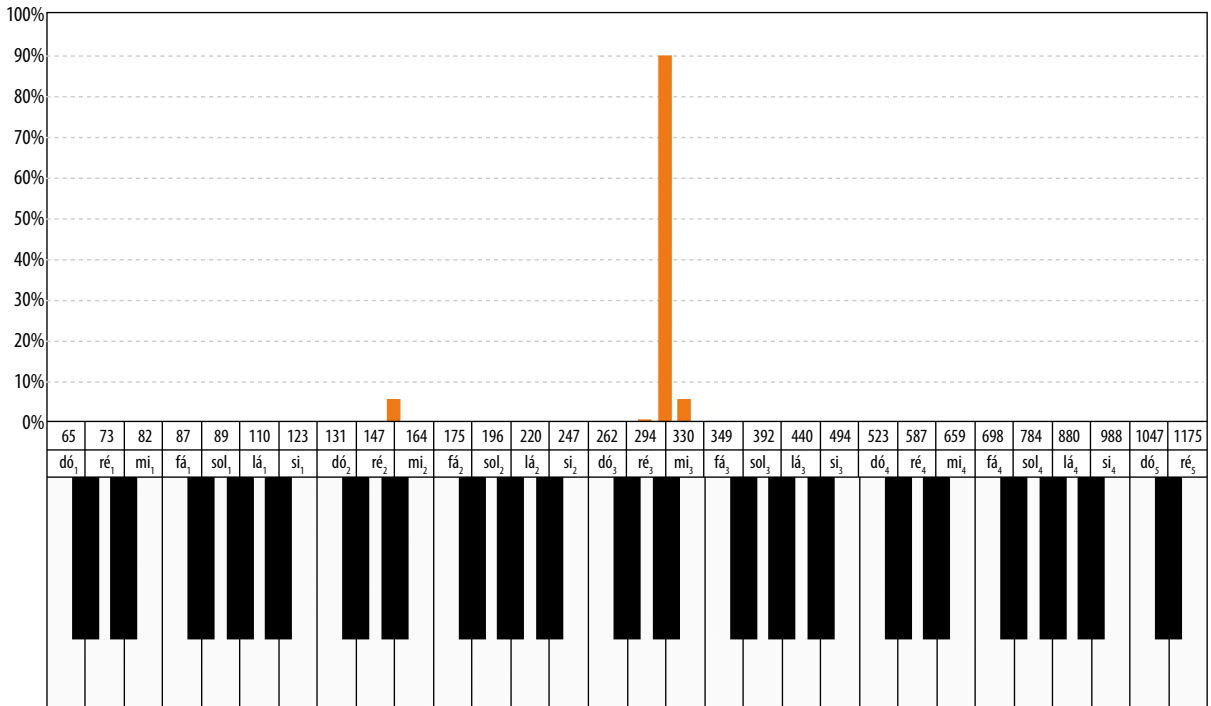
Pitch-matching

Pitch-matching was administered individually to the participants in a quiet environment, with sound stimuli presented under free-field conditions at a self-selected habitual loudness. In Task 1, the individual had to listen to an isolated musical tone (Chart 1) and then immediately repeat it vocally, a task repeated with five different tones. In Task 2, the individual had to listen to a 3-tone sequence (Chart 1) and then immediately repeat the sequence vocally, a task again repeated using five different sequences. The vocal reproductions were directly captured into a portable computer by means of a head-mounted microphone that

had a flat frequency response; it was placed at 45° and 5 cm away from the mouth of the participant. The samples were recorded using Sound Forge software version 4.5c and imported into Vocalgrama 1.8i (CTS Informática). The parameters used are shown in Chart 1.

Data analysis. All voice samples were subjected to computerized acoustic analysis by means of the Vocalgrama software. Vocalgrama uses autocorrelation to determine F0; a filter to remove noise, available on the software, was used to reduce artifacts. The frequency of an individual's vocal imitation was compared to the frequency of the original tone. A correct match was considered to be when the reproduction had the same fundamental frequency as the original to within a semitone (Figure 1), and the individual was then considered to have accurate pitch-matching. In cases where the vocal imitation and the original tone had different frequencies, the match was considered wrong (Figure 2). Participants who were able to sing correct sequences of tones with 100% accuracy were considered as able to sing in-tune. However, if a participant was unable to correctly repeat the sequences they were considered as singing off-tune [3]. The fundamental frequency extraction was performed outline.

Chart 1. Selected stimuli for the pitch-matching scanning

Task	Order of stimuli		Men	Women
Task 1	Rrst stimulus		E^3	E^4
	Second stimulus		$G\#^3$	$G\#^4$
	Third stimulus		F^3	F^4
	Fourth stimulus		B^3	B^4
	Fifth stimulus		$D\#^3$	$D\#^4$
Task 2	First sequence	First stimulus	E^3	E^4
		Second stimulus	A^3	A^4
		Third stimulus	$F\#^3$	$F\#^4$
	Second sequence	First stimulus	G^3	G^4
		Second stimulus	$D\#^3$	$D\#^4$
		Third stimulus	F^3	F^4
	Third sequence	First stimulus	D^3	D^4
		Second stimulus	G^3	G^4
		Third stimulus	E^3	E^4
	Fourth sequence	First stimulus	D^3	D^4
		Second stimulus	F^3	F^4
		Third stimulus	A^3	A^4
	Fifth sequence	First stimulus	$A\#^3$	$A\#^4$
		Second stimulus	$F\#^3$	$F\#^4$
		Third stimulus	E^3	E^4

Figure 1. Example of correct tuning in the computerised pitch-matching system. Here the tone produced was D#4 (gray key) and the corresponding peak of the frequency histogram shows that the note was reproduced vocally at the same frequency. (Source: Vocalgrama 1.8i, CTS Informática)
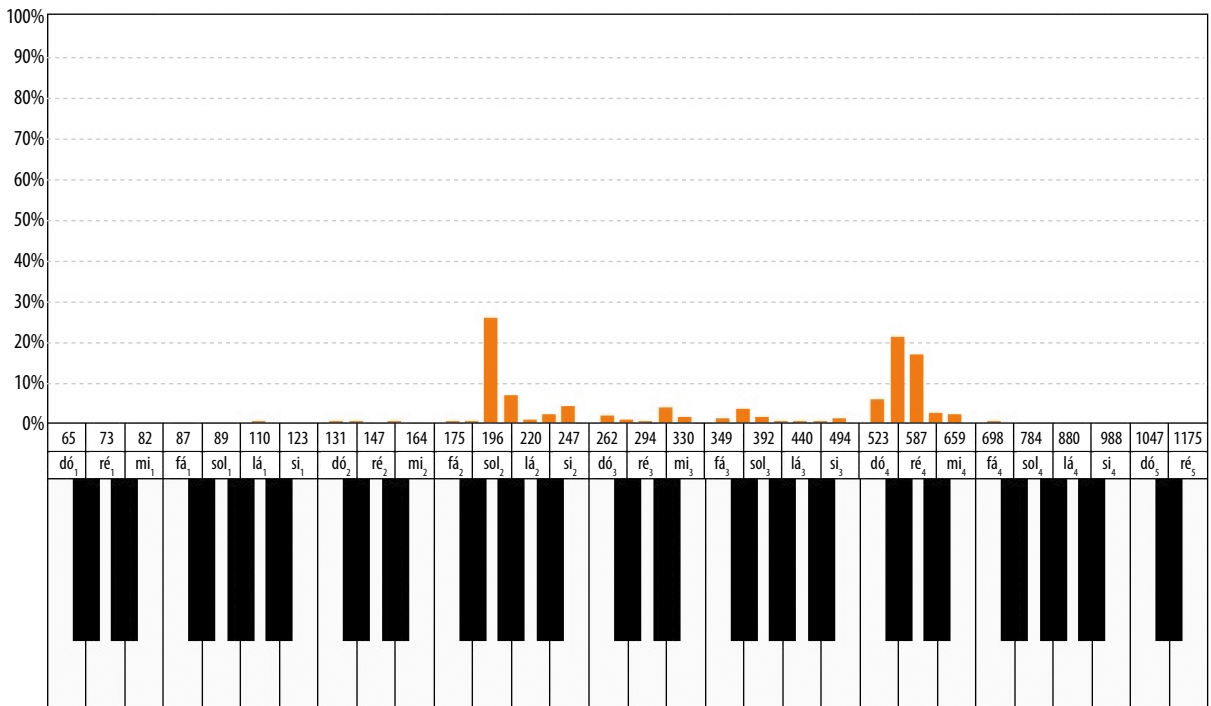
Figure 2. Example of incorrect tuning in the computerised pitch-matching system. Here the tone produced was again D#4 (gray key) but the peak of the frequency histogram of the vocal reproduction occurred at G4, 4 semitones away from the original tone. (Source: Vocalgrama 1.8i, CTS Informática)

Chart 2. Acquisition parameters of the FFR system. Key: RE, right ear; LE, left ear; ms, milliseconds; ABR, auditory brainstem response; FFR, frequency following response

PARAMETER	Click-ABR	FFR
Equipment	Biologic Navigator Pro	Biologic Navigator Pro
Stimulated ear	RE/LE	RE/LE
Type of stimulus	Click	Speech
Duration of stimulus	0.1 ms	40 ms
Polarity of stimulus	Rarefaction	Alternating
Intensity of stimulus	80 dB nHL	80 dB
Rate of stimuli	19.3/s	10.9/s
Number of scans	2000	3000
Replicability	2 collections of 2000 stimuli	2 collections of 3000 stimuli
Filter	10–1500 Hz	100–2000 Hz
Window	10.66 ms	85.33 ms
Transducer	Insert (ER-3A; Natus Medical)	Insert (ER-3A; Natus Medical)

Electrophysiological evaluation

The electrophysiological evaluation was conducted using the Biologic Navigator Pro equipment (Natus, USA) in an acoustically prepared soundproof and electrically shielded room. Subjects were seated in a reclining chair in a comfortable position. The skin of the subject's scalp was cleaned with abrasive paste before fixing the electrodes in place with conductive paste and adhesive tape. Impedance was kept below 3 kΩ and the inter-electrode impedance was less than 2 kΩ. The electrodes were positioned according to the 10-20 system, e.g active electrode at the apex (Cz), reference electrode on the ipsilateral mastoid, and ground electrode on the contralateral mastoid [26]. The parameters used are shown in Chart 2.

During testing, subjects were instructed to keep their eyes closed in order to avoid artifacts. If necessary, changes were made to the position of the subject in order to guarantee stable recording conditions. Runs containing more than 10% artifacts were repeated.

All analyses were performed offline and the response waveforms were visually identified and marked manually by an audiologist who was blinded to each participant's age, gender, and group (CG or EG).

The ABR responses were recorded on the right and left ears separately at 80 dBnHL. Two waveforms were collected to verify reproducibility. The presence and absolute latencies of waves I, III, and V to 80 dBnHL were analysed, as well as the interpeak intervals I–III, III–V, and I–V according to the normality criteria proposed by the NavPro Biologic system.

The FFR responses were recorded via monaural stimulation of the right ear. The choice of this stimulation mode is related to the right ear advantage and the dominance of the left side of the brain for language skills [27-30].

The analysis was performed in the time domain. Latency and amplitude values of the seven waves (V, A, C, D, E, F, and O) elicited by the syllable /da/ were based on the analysis criteria of previous published studies [14, 24-29]. The analysis focused on four major elements: (i) the onset stimulus portion (represented by waves V and A); (ii) the transition period between the consonant and vowel (wave C); (iii) the sustained portion (represented by waves D, E, F); and (iv) the offset portion, represented by wave O (Figure 3). Subsequently, the latency and amplitude values were analysed for each identified wave. In case a wave was not detected, the wave was described as being absent and the data for this wave was not analysed. In addition, analysis of the VA complex was also carried out, which involved:

(i) the slope of the VA complex (µV/ms), which is related to the temporal synchronization of the response generators [27];

(ii) area of the VA complex (µV × ms) which is related to the amount of activity that contributes to generation of the wave [27].
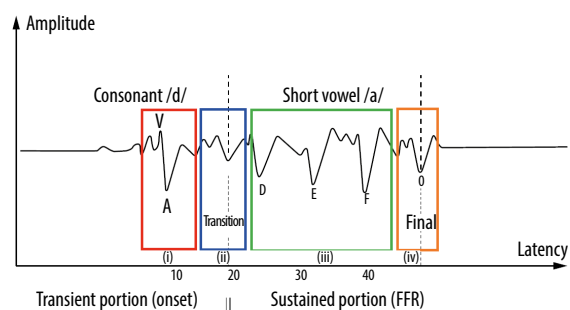
Figure 3. Schematic of FFR responses to synthesised syllable /da/. Analysis was done on the four major elements labelled

Statistical analysis

In order to compare groups (in-tune and off-tune) across each wave, an ANOVA procedure was used, testing for the age and gender variables as well as their interactions. The variables group, gender, and age were fixed at two levels each. The ANOVA used the Fisher–Snedecor distribution to determine whether there was a significant difference among groups or their interactions. The FFR response presents seven wave peaks, which are also related to each other, so the *p*-values were calculated from the ANOVA analysis. In order to test the homogeneity of the sample, a Pearson chi-square test was applied. The level of significance was set at 5% ($p \leq 0.05$). Statistical analyses were conducted using R-project (www.r-project.org).

Results

Homogeneity of the sample

Tables 1 and 2 give statistical descriptions of the demographic data of the in-tune and off-tune individuals based on the variables age and gender. They show that there is homogeneity between the two groups regarding age and gender.

Frequency Following Response

Table 3 displays a comparison between the in-tune and off-tune individuals in terms of FFR latency for the right ear. It shows statistically significant latency differences for waves A ($p = 0.045$), C ($p = 0.002$), D ($p = 0.030$), and F ($p = 0.046$) for the in-tune and off-tune groups.

Table 4 displays the comparison between in-tune and off-tune individual in terms of FFR amplitude (μV) in the right ear. There were no statistically significant differences between the groups.

Discussion

Homogeneity of the sample: Our groups were statistically homogeneous, both in terms of gender and age.

Frequency Following Response: The FFR is an electrophysiological examination that enables one to localize and monitor speech sound processing in the subcortical area. The syllable /da/, which is used in the FFR test reflects both phonetic/filter (transient) and prosodic/source (periodic) acoustic features with remarkable fidelity [31-33]. Surveys show that, due to neural plasticity, musicians show an advantage in this coding derived from cortical activation arising from extensive musical training and practice. However, it is not known how the encoding of verbal stimuli in off-tune individuals occurs. The aim of this study was to investigate properties of verbal sound coding in the subcortical area of off-tune individuals, and a simple way to explore this aspect was to make a comparison with a group of in-tune individuals.

Latencies values

(i) *Onset portion (consonant):* (i.a) *Waves V and A:* Of these two waves, there was a statistically significant difference in the latency values for wave A only ($p = 0.045$). This study

presented the sound stimulus /da/, and wave A is the onset portion of this stimulus comprising responses to encoding of the consonant /d/. The most robust responses with early latencies values were observed with in-tune individuals, which suggests that this group had an effective synchronous neural response from the beginning of the sound stimulus, indicating high functionality of the peripheral auditory system [8, 35].

The results of this study demonstrate that off-tune individuals differ in their responses compared to tuned individuals in terms of their initial sound processing. They show a significant delay in the latency of wave A.

Previous research has shown changes for latency of wave A in patients with attention deficit hyperactivity disorder (ADHD), epilepsy, dyslexia, hearing loss, scholastic difficulties, auditory processing disorder, and/or learning disabilities. These changes were associated with faulty processing of speech in the auditory system or by inaccuracies in the inferior colliculus, which is probably the neuronal generator of wave A [7, 12, 15, 22, 27, 32, 36].

(i.b) *VA Slope and VA Area:* There were no statistically significant differences in the values of the VA complex (VA Slope and VA Area) between the groups. VA Slope is related to the synchronization between the response generators, while VA Area is related to the amount of activity that contributes to wave generation.

(ii) *Transition period between the consonant and vowel*: There were statistically significant differences in the values of the wave C ($p = 0.002$). Wave C marks the transition zone between consonants and vowels, and showed worse latency values in off-tune individuals compared to in-tune individuals. There are reports of wave C delays in some pathology which might be due to a decrease in neural firings that would cause changes in processing time [34, 37-39]. The correct coding of a sound stimulus after a short time is important in precisely discriminating a message. The processing of wave C depends on a good temporal resolution skill that seems to be impaired in off-tune individuals.

(iii) *Sustained portion (vowel):* There were statistically significant differences in the values of the wave D ($p = 0.003$) and wave F ($p=0.046$). The coding of fundamental frequency as well as the harmonic structure of speech originates in the adult midbrain [40, 41]. The neural generators of FFRs are modified by sensory, limbic, and cognitive circuits [42]. The sustained portion is comprised of waves D, E, and F, which correspond to the fundamental frequency (F0), and is associated with particular sound features that may contribute to understanding a message, such as pitch, phonemic identity, and temporal cues [27, 43, 44]. In the case of off-tune individuals, who are the focus of this study, there was a particular response pattern involving a significant increase in latency of waves D and F compared to in-tune individuals.

The latency changes found for waves D and F can be explained by a lack of higher cognitive processing ability in off-tune individuals. Recent studies have shown that linguistic and/or musical processing trigger an extensive bilateral neuronal network in the auditory cortex [45]. Thus, musicians tend to

Table 1. Statistical description of demographic data based on the variable age

Variable	Group	N	Mean	SD	Minimum	Median	Maximum	*p*-value
	Off-tune	20	30.1	11.0	20.0	25.5	57.0	
Age	In-tune	17	35.4	5.7	28.0	34.0	49.0	0.07
	Total	37	32.1	10.9	20.0	33.0	57.0	

Key: SD, standard deviation. Fisher's exact test for count data

Table 2. Statistical description of demographic data based on the variable gender

Variable	Sex	Group				*p*-value
		In-tune		Off-tune		
		Number (*N*)	Percentage (%)	Number (*N*)	Percentage (%)	
Gender	Male	7	41.2	10	50	0.84
	Female	10	58.8	10	50	

Chi-square Pearson test

Table 3. Comparison between in-tune and off-tune individuals in terms of FFR latency (ms) in the right ear

	IN-TUNE			OFF-TUNE				
Component	Mean	Median	SD	Mean	Median	SD	Diff in Mean	*p*-value
V	6.56	6.62	0.28	6.67	6.70	0.30	−0.11	0.311
A	7.54	7.62	0.31	7.79	7.87	0.38	−0.25	0.045*
C	18.06	18.12	0.51	18.77	18.62	0.72	−0.72	0.002*
D	22.21	22.20	0.63	22.94	22.58	1.25	−0.74	0.030*
E	30.80	30.70	0.30	31.83	31.03	2.27	−1.03	0.065
F	39.15	39.20	0.30	40.22	39.36	2.19	−1.07	0.046*
O	47.94	47.95	0.58	49.33	48.20	3.32	−1.39	0.080
Slope	0.28	0.26	0.07	0.23	0.24	0.10	0.04	0.133
Area	0.26	0.25	0.09	0.27	0.26	0.10	0.00	0.830

Key: SD, standard deviation. F-test from ANOVA *p*-value definition. *p-value < 0.05

Table 4. Comparison between in-tune and off-tune individuals in terms of FFR amplitude (μV) in the right ear

	IN-TUNE			OFF-TUNE				
Component	Mean	Median	SD	Mean	Median	SD	Diff in Mean	*p*-value
V	0.10	0.09	0.04	0.08	0.08	0.06	0.02	0.198
A	0.16	0.18	0.05	0.16	0.16	0.05	0.00	0.981
C	0.09	0.07	0.09	0.08	0.05	0.09	0.01	0.830
D	0.07	0.06	0.04	0.11	0.08	0.08	0.04	0.087
E	0.17	0.18	0.05	0.18	0.17	0.07	-0.01	0.754
F	0.18	0.16	0.08	0.14	0.12	0.11	0.03	0.391
O	0.11	0.10	0.06	0.17	0.12	0.23	-0.06	0.334

Key: SD, standard deviation. F-test from ANOVA *p*-value definition

have stronger mental imagery and higher-order music cognition, occupying higher tract volumes, which suggest that their training has led to improvements in neural circuitry and transmission speed of auditory information [45, 46]. Musical training involves years of sensory-motor training and improves the ability to attend to the fine-grained features of musical sounds, including pitch, timing, and timbre [47]. This portion of the FFR might therefore represent responses related to the pitch and timbre of the sound stimulus, and which are essential to successful auditory perception. Since in music there are fairly direct parallels between the pitch and harmonic features of a vowel, off-tune individuals may be deficient in this area.

(iv) *Offset portion:* There were no statistically significant differences in the latency and amplitude values of wave O between the groups.

Amplitude values

The amplitude measurements do not seem to be a suitable parameter for identifying changes in speech coding between the in-tune and off-tune groups. Both groups showed high variability for all waves analysed (V, A, C, D, E, F, and O). These findings corroborate previous FFR studies which also highlight that the amplitude measures are not very reliable in distinguishing between normal and pathological individuals [34].

FFR responses of the off-tune group

This study revealed a difference between in-tune and off-tune individuals in neural processing of the syllable /da/. A possible explanation is that an individual with a good knowledge and experience of music is more likely to have developed efficient processing of language [48, 49]. The number of research studies to understand speech processing in musical individuals is increasing. Research on music neurobiology is a case in point, particularly on how the brain stem functions. The brain stem seemed to have a passive role in speech processing but now stands out as taking an active role in the neural decoding of sounds. Significantly, it seems to have the possibility of being continually modified with experience and stimulation throughout life.

The multisensory nature of music may have an impact on vocal production by involving motor, auditory, and vocal mechanisms. Research has shown that musicians have less harmonic jitter in their voices, either when singing or talking, and this suggests that vocal training has an effect on vocal tract resonances during speech [50]. Failures in vocal perception may indicate processing, memory, language, or speech production problems.

FFR testing provides a standard metric of auditory encoding and might be useful in disentangling the roles of pitch, timbre, and timing in normal and pathological individuals [9]. Moreover, it might be used to monitor the progress of speech therapy interventions for improving vocal tuning, since there is a strong correlation between auditory perception and speech production. Thus, a new method for understanding how CANS can affect voice quality, auditory processing, and speech perception might become available. It seems that in certain individuals the responses could be stable, interpretable, and meaningful [31, 28, 29, 51].

Off-tune individuals cannot monitor their own hearing and vocal ability. The FFR has already proven to be an objective and effective instrument for identifying changes in sound perception, and our study shows it can also allow a differential diagnosis to be made between in-tune and off-tune people.

The FFR responses show how neural coding may vary from one population to another. Listeners with a given pathology have a specific pattern of response that may help explain how neural processes occur in the case of ordinary communication [33]. In the case of off-tune individuals, there is a particular pattern characterized by an increase in the latencies of waves A, C, D, and F when compared with in-tune individuals. Previous studies have shown that waves C and F have the most stable peaks [28, 29]. Interestingly, the in-tune individuals had a smaller latency (higher sound processing speed) of all FFR waves compared to normal subjects. It seems they have an enhanced perception of spectro-temporal features, particularly rapid changes and harmonic patterns.

Limitation of this study

The present study has some limitations. FFR responses can be analyzed in either the time or frequency domain, and in the present study it was analyzed in the time domain only. It is also important to conduct further research with a larger number of individuals to be able to understand and compare the findings seen in this study.

Conclusion

The results of this research demonstrate that the neurophysiological responses of the Frequency Following Response can be affected by tuning ability. The group of off-tune individuals had prolonged A, C, D, and E waves. This research provides support for differential diagnosis based on the FFR and for its use in therapeutic intervention of vocal tuning ability.

References

1. Kimura D. Cerebral dominance and the perception of verbal stimuli. Can J Psychol, 1961;15(3):166-71.
2. Asaridou SS, McQueen JM. Speech and music shape the listening brain: evidence for shared domain-general mechanisms. Front Psychol, 2013;4:321.
3. Moreti F, Pereira LD, Gielow I. Pitch-matching scanning: comparison of musicians and non-musicians' performance. J Soc Bras Fonoaudiol, 2012;24(4):368-73.
4. Mawhinney T. Tone-deafness and low musical abilities: an investigation of prevalence, characteristics and tractability. Kingston: Queen's University; 1986.

5. Heresniak M. The care and training of adult bluebirds: teaching the singing impaired. J Singing, 2004;61(1):9-25.

6. Ishii C, Arashiro P, Pereira L. Ordering and temporal resolution in professional singers and in well tuned and out of tune amateur singers. Pro Fono, 2006;18(3):285-92.

7. Strait D, Parbery-Clark A, Hittner E, Kraus N. Musical training during early childhood enhances the neural encoding of speech in noise. Brain Lang, 2012;123(3):191-201.

8. Musacchia G, Strait D, Kraus N. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear Res, 2008;241(1-2):34-42.

9. Strait D, Parbery-Clark A, O'Connell S, Kraus N. Biological impact of preschool music classes on processing speech in noise. Develop Cog Neurosci, 2013;6:51-60.

10. Nascimento F, Monteiro R, Soares C, Ferreira M. Temporal sequencing abilities in musicians violinists and non-musicians. Arq Int Otorrinolaringol, 2010;14(2):217-24.

11. Strait DL, O'Connell S, Parbery-Clark A, Kraus N. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. Cereb Cortex, 2014;24(9):2512-21.

12. Wible B, Nicol T, Kraus N. Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. Biolog Psychol, 2004;67(3):299-317.

13. Kishon-Rabin L, Amir O, Vexler Y, Zaltz Y. Pitch discrimination: are professional musicians better than non-musicians? J Basic Clin Physiol Pharmacol, 2001;12(2):125-43.

14. Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing: implications for language and music. Ann N Y Acad Sci, 2009;1169:543-57.

15. Sanfins M, Borges L, Ubiali T, Colella-Santos M. Speech-evoked auditory brainstem response in the differential diagnosis of scholastic difficulties. Braz J Otorhinolaryngol, 2015.

16. Sanfins M, Colella-Santos M. A review of the clinical applicability of speech-evoked auditory brainstem responses. J Hear Sci, 2016;6 (1):9-16.

17 Blumstein S, Stevens K. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. J Acoust Soc Am, 1979;66(4):1001-17.

18. Ladefoged P. A Course in Phonetics. Boston: Thomson Higher Learning; 2006.

19. Bachorowski J-A, Owren M. Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. Psychol Sci, 1995;6(4):219-24.

20. Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. Proc Natl Acad Sci U S A, 2009;106(31):13022-7.

21. Dhar S, Abel R, Hornickel J, Nicol T, Skoe E, Zhao W, et al. Exploring the relationship between physiological measures of cochlear and brainstem function. Clin Neurophysiol, 2009;120(5):959-66.

22. Basu M, Krishnan A, Weber-Fox C. Brainstem correlates of temporal auditory processing in children with specific language impairment. Dev Sci, 2010;13(1):77-91.

23. Rana B, Barman A. Correlation between speech-evoked auditory brainstem responses and transient evoked otoacoustic emissions. J Laryngol Otol, 2011;125(9):911-6.

24. Davis H, Silverman RS. Hearing and Deafness. New York: Rinehart & Winston; 1970

25. Jerger J. Clinical experience with impedance audiometry. Arch Otolaryngol, 1970; 92: 311-24.

26. Jasper HH. The ten–twenty system of the International Federation. Electroencephalogr Clin Neurophysiol, 1958; 10: 371-75.

27. Skoe E, Kraus N. Auditory brainstem response to complex sounds: a tutorial. Ear Hear, 2010; 31:320-24.

28. Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clin Neurophysiol, 2004;115:2021-30.

29. Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N. Auditory training improves neural timing in the human brainstem. Behav Brain Res, 2005;156(1):95-103.

30. Keith R. Dichotic listening in children. In: Beasley D, editor. San Diego: College-Hill Press; 1984.

31. Song JH, Nicol T, Kraus N. Test–retest reliability of the speech-evoked auditory brainstem response. Clin Neurophysiol, 2011;122(2):346-55.

32. Johnson KL, Nicol TG, Kraus N. Brain stem response to speech: a biological marker of auditory processing. Ear Hear, 2005;26(5):424-34.

33. Sinha SK, Basavaraj V. Speech evoked auditory brainstem responses: a new tool to study brainstem encoding of speech sounds. Indian J Otolaryngol Head Neck Surgery, 2010;62(4):395-9.

34. Sanfins MD, Borges LR, Ubiali T, Donadon C, Hein TAD, Hatzopoulos S, et al. Speech-evoked brainstem response in normal adolescent and children speakers of Brazilian Portuguese. Int J Pediatr Otorhinolaryngol, 2016:12-9.

35. Wong P, Skoe E, Russo N, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci, 2007;10 (4):420-2.

36. Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiol, 2010;47(2):236-46.

37. Gabr TA, Darwish ME. Speech auditory brainstem response audiometry in children with specific language impairment. Hear Bal Commun, 2016;14(1):50-7.

38. King C, Warrier C, Hayes E, N K. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett, 2002;319:111–5.

39. Shennawy A, Khosht M, Ghannoum H, Meguid N. Electrophysiologic assessment of auditory function in children with autism and attention deficit and hyperactivity disorder. J Hear Sci, 2014;4(3):26-34.

40. Krishnan A. Human frequency-following responses: representation of steadystate synthetic vowels. Hear Res, 2002;166:192-201.

41. Galbraith G. Two-channel brainstem frequency-following responses to pure tone and missing fundamental stimuli. Electroencephalogr Clin Neurophysiol, 1994;92:321-30.

42. White-Schwoch T, Davies EC, Thompson EC, Woodruff Carr K, Nicol T, Bradlow AR, et al. Auditory-neurophysiological responses to speech during early childhood: effects of background noise. Hear Res, 2015;328:34-47.

43. Krishnan A, Gandour J, Ananthakrishnan S, Bidelman G, CJ S. Functional ear (a)symmetry in brainstem neural activity relevant to encoding of voice pitch: a precursor for hemispheric specialization? Brain Lang, 2011;119:226-31.

44. Gockel H, Carlyon R, Mehta A, Plack C. The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. JARO, 2011;12:767-82.

45. Schneider P, Scherg M, Dosch H, Specht H, Gutschalk A, Rupp A. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci, 2002;5:688–94.

46. James CE, Oechslin MS, Van De Ville D, Hauert CA, Descloux C, Lazeyras F. Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. Brain Struct Funct, 2014;219(1):353-66.

47. Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nature Rev Neurosci, 2010;11:599-605.

48. Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci, 2009;29(45):14100-7.

49. Strait D, Kraus N, Parbery-Clark A. Musical experience shapes top-down mechanisms: evidence from masking and auditory attention performance. Hear Res, 2010;261:22-9.

50. Wolfe J, Garnier M, Smith J. Vocal tract resonances in speech, singing, and playing musical instruments. HFSP Journal, 2008;3(1):6-23.

51. Hornickel J, Knowles E, Kraus N. Test–retest consistency of speech-evoked auditory brainstem responses in typically-developing children. Hear Res, 2012;284(1-2):52-8.