Brain & Language 260 (2025) 105506

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Revisiting the 40-Hz gamma response: Phase-locked neural activity along the human auditory pathway relates to bilingual experience

Luan Tonelli [a], Parker Tichko [b], Erika Skoe [c,*]

[a] Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, CT 06269, USA
[b] Department of Psychological Sciences, Developmental Psychology Program, University of Connecticut, Storrs, CT 06269, USA
[c] Connecticut Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, CT 06269, USA

ARTICLE INFO

Keywords:
Frequency-following response
Bilingualism
Plasticity
Auditory system
Auditory
Computational model

ABSTRACT

Spoken language experience influences brain responses to sound, but it is unclear whether this neuroplasticity is limited to speech frequencies (>100 Hz) or also affects lower gamma ranges (~30–60 Hz). Using the frequency-following response (FFR), a far-field phase-locked response to sound, we explore whether bilingualism influences the location of the strongest response in the gamma range. Our results indicate that the strongest gamma response for bilinguals is most often at 43 Hz, compared to 51 Hz for monolinguals. Using a computational model, we show how this group difference could result from differential subcortical activation. These results shed light on the well-known but under-explored variability observed in the gamma range and highlight that FFRs are a composite of neural activity from both subcortical and cortical sources. Additionally, our findings emphasize that individual auditory experiences can uniquely shape subcortical activation, influencing FFRs below speech frequencies.

1. Introduction

Far-field neurophysiological responses to sound, such as the frequency-following response (FFR), provide a noninvasive window for studying how individual differences in abilities and experience manifest in the central auditory nervous system (CANS) (Skoe & Kraus, 2010, Kraus and White-Schwoch, 2015, Nozaradan et al., 2016, White-Schwoch et al., 2017). FFRs were first recorded noninvasively in humans in the early 1970s (Moushegian et al., 1973). Today, FFRs are used for a variety of purposes in human populations, including studying individual differences arising from auditory experience (e.g., experience with music, short-term training) or relating to language exposure and ability (Krishnan et al., 2005, Musacchia et al., 2007, Coffey et al., 2016, Bidelman et al., 2018, Nozaradan et al., 2016, Kraus and Nicol, 2017).

FFRs are phase-locked neurophysiological responses recorded from the scalp that reflect the periodicity of an acoustic stimulus. Synthetic (e.g., sinusoids, complex tones) and ecological sounds (e.g., speech and music) can be used to elicit the FFR and study its neural sources (Skoe and Kraus, 2010, Coffey et al., 2019, Lerud et al., 2023). Early studies using experimental animal models (Gardi et al., 1979, Snyder and Schreiner, 1984) indicated that the FFR to frequencies in the vocal pitch (from ~ 85 to 400 Hz) has subcortical sources. This led to a similar FFR interpretation for humans, namely that (a) cortical areas dominate the generation of FFRs for lower stimulus frequencies (<80 Hz) and (b) subcortical auditory nuclei, although responsive to lower frequencies, dominate the scalp-recorded response to higher frequencies in the speech range (>150 Hz; Galambos et al., 1981, Makela and Hari, 1987, Ribary et al., 1991, Kuwada et al., 2002), with phase-locking becoming progressively weaker to non-existent above 1000 Hz (Batra et al., 1986, Clinard et al., 2010). More modern studies (Kraus & White-Schwoch, 2015, Coffey et al., 2019, Skoe et al., 2017, Tichko and Skoe, 2017, Bidelman, 2018, Lerud et al., 2023) converge to suggest that the scalp-recorded phase-locked response to any given stimulus is a far-field aggregation of activity from multiple generators in auditory neuroaxis, with the degree of contribution from any one source being dictated by the stimulus frequency, the recording methods (e.g., electrode montages, EEG vs. MEG), and potentially also the demographic and ecological experiences of the participant and their attentional state (Bidelman et al., 2018; Coffey et al., 2019; Hartmann et al., 2019). These more modern accounts of FFR sources to speech frequencies, including our recent work (Lerud et al., 2023), allow for the possibility that the auditory cortex contributes – albeit weakly – to phase-locked responses recorded at the scalp in response to speech frequencies.

In addition to reflecting multiple sources, the FFR, even to simple

---

* Corresponding author.
E-mail address: erika.skoe@uconn.edu (E. Skoe).

https://doi.org/10.1016/j.bandl.2024.105506
Received 4 October 2023; Received in revised form 4 April 2024; Accepted 27 November 2024
Available online 13 December 2024

acoustic stimuli, is a multidimensional response that can be analyzed with respect to its latency, amplitude, phase, and spectral profile (e.g., location of spectral peaks and valleys, among other approaches (Skoe & Kraus 2010, Krizman & Kraus, 2019). FFR spectral amplitude varies as a function of the stimulus's fundamental frequency (F0) in a predictable yet nonlinear fashion. The waxing and waning of the FFR amplitude is a well-known phenomenon that Gardi et al., 1979 described more than 40 years ago as a "characteristic" of the FFR. These characteristic frequency-dependent amplitude fluctuations become particularly apparent when a wide range of stimulus frequencies is tested (Kuwada et al., 2002, Skoe and Tichko, 2017, Gransier et al., 2020). Our recent paper (Tichko & Skoe, 2017) reported FFRs to more than 90 frequencies, ranging from 16.35 to 880 Hz. In our analysis of the FFR spectral amplitude at the stimulus F0, we observed that certain frequency ranges consistently elicited robust responses, which we have termed the "FFR local maxima". Notably, the frequency that produced the overall largest response (i.e., the global maximum) was in the 40–60 Hz range – this agrees with earlier findings showing a response maximum in this gamma range (Hari et al., 1989, Stapells et al., 1984, Lee et al., 1984, Galambos et al., 1981). These earlier works coined this the "40-Hz response" (Hereafter, we refer to this phenomenon as the gamma range max or GRmax, and we use this term interchangeably with the "40-Hz response").

In our earlier work (Tichko and Skoe, 2017), we argued that the waxing and waning in amplitude across stimulus frequency indicate that the FFR reflects multiple generators, similar to what Kuwada et al. proposed in 2002. Using a computational model of FFR generation based on Gardi et al. (1979) and informed work by Kuwada et al., 2022, we then demonstrated how this waxing and waning pattern of the FFR amplitude across stimulus frequencies could arise from multiple sources responding either in-phase or out-of-phase, respectively (Fig. 2**-A and B**). Although we tested only three healthy young adult participants in this earlier work, all showed the same general frequency-dependent "footprint" (i.e., consistency in the general location of the FFR maxima and minima) and the presence of the "40-Hz" response (i.e., increased FFRs near 40-Hz), characteristics of the FFR that align with what had been seen in earlier studies (Gardi et al., 1979, Kuwada et al. 2022). Individual differences were, however, still evident in the exact stimulus frequency that produced the GRmax. We proposed that, as one possible explanation, these differences in the location of the global maximum could arise from individual variations in auditory experience.

Consistent with our prior study, a recent study (Gransier et al., 2021) showed that phase-locked auditory responses between 0.5 to 100 Hz have significant intersubject variability using a technique similar to FFR, the Auditory Steady-State Response (ASSR). Like our earlier work, the Gransier et al. study used a large set of stimuli with small frequency spacing (Kuwada et al., 2002); this contrasts with more conventional studies that use only a single or a small set of stimulus (Galbraith, 1994, Gorina-Careta et al., 2021). Although all participants had their ASSR maximum near 40 Hz in the Gransier et al., 2021, the exact stimulus frequency that produced the maximum phase-locked response varied across participants and fell between 30–60 Hz. The authors did not elaborate on the demographic characteristics of the participants or consider that differences in auditory experience could underlie this variation. However, they did consider the neuroanatomic sources of this variation using source analysis applied to high-density EEG recordings. They concluded that intersubject variability is greater for phase-locked responses originating from cortical than subcortical regions. For the "40-Hz response", auditory cortex is considered the dominant source for ASSRs and FFRs; however, thalamus and brainstem are also involved in generating the scalp-recorded phase-locked response but are comparatively weaker (Luke et al., 2017, Farahani et al., 2017, Herdman et al., 2002). Indeed, studies in experimental animals suggest when inferior colliculus is lesioned, activity in the gamma range decreases by 40 % (Tsuzuku, 1993). In theory, individual differences in the level of subcortical activation could also influence the location of the GRmax, given the non-negligible role of subcortical structures in generating phase-locked responses, regardless of the stimulus frequency (Bidelman 2018).

The "40-Hz response" is an evoked response potential that is well-described in the auditory literature and shows parallels in other sensory domains, including the visual system (Galambos et al., 1981, Artieda et al., 2004, McDermott et al., 2018). Previous auditory studies have consistently demonstrated strong neural phase-locked activity for tonal stimuli within this lower gamma frequency range (i.e., 30–60 Hz) (Picton et al., 2003, Purcell et al., 2004, Poulsen 2007). Other evidence for individual differences in this gamma frequency range can also be found. For instance, Zaehle et al. (2010), in a dataset of 20 healthy females, found that the ASSR maximum in the gamma range (i.e., the GRmax) varied across individuals, from ~ 30–55 Hz. Baltus and Herrmann (2015) showed that listeners whose GRmax was located at a higher stimulus frequency had better auditory processing abilities. Specifically, those with a higher GRmax frequency had lower gap detection thresholds, suggestive of better auditory temporal acuity. Moreover, Poulsen et al. (2007) demonstrated a positive relationship between the GRmax frequency and age, indicating that the frequency eliciting the frequency that produces the strongest phase-locked response increases with age.

Here, in a data-driven analysis, we explore the role of bilingual language experience on the GRmax phenomenon using an FFR dataset that includes English monolinguals and bilinguals, i.e., individuals with substantial exposure to two spoken languages. To our knowledge, this is the first study to examine this phenomenon in bilinguals. While there is no literature to directly draw on to make predictions with respect to bilingualism, the 40-Hz response has been connected to a broad spectrum of perceptual and cognitive functions within neuropsychiatric disorders, such as attention, memory, and sensory integration, functions that are undoubtedly vital for bilingual individuals navigating multiple language systems (Griškova-Bulanova et al., 2022, Parciauskaite et al., 2019). The 40-Hz response is also being investigated for its potential to explain individual differences in cognitive abilities like inhibitory control (see review – Declerck & Koch, 2023, Rossi et al., 2018, Rodriguez-Fornells et al., 2006). This research is directly relevant to the current work, as inhibitory control is proposed to be enhanced in bilinguals (Declerck & Koch, 2023, Rossi et al., 2018, Blumenfeld & Marian, 2011, Fornells et al., 2006). Additionally, studies of language-impaired populations suggest that these populations often display altered gamma-band activity, suggesting a fundamental role of this neural phenomenon in language processing and cognitive functions. Altered gamma-band responses in language-impaired individuals may reflect disruptions in neural synchronization and information processing, which are critical for effective language comprehension and production (Uhlhaas & Singer 2012, Pulvermüller et al., 2001). This body of work underscores the importance of gamma frequency range activities in understanding the neural mechanisms behind cognitive functions and their potential implications for bilingual individuals.

While the 40-Hz response has not been studied in bilinguals, other FFR dimensions have been linked to auditory exposure history including music and language learning (Anderson et al., 2018; Bones, Hopkins, Krishnan, & Plack, 2014; Krishnan, Xu, Gandour, & Cariani, 2005; Tichko & Skoe, 2017; Valle Rodrigues et al., 2019). Studies show that individuals with extensive exposure to two spoken languages (e.g., bilinguals) have stronger FFRs to speech stimuli with F0s in the ~ 100 Hz range (Krizman et al., 2012; Skoe et al., 2017; Skoe et al., 2022). However, FFR differences between bilinguals and monolinguals for stimulus frequencies outside this F0 range, including the gamma range, have not been studied. There is also a knowledge gap in understanding where along the auditory neuroaxis differences between monolinguals and bilinguals first emerge, although previous work implicates the auditory brainstem (Krizman et al., 2012, Krizman et al., 2015, Giroud et al., 2020). To help address this gap, we recorded FFRs to auditory stimuli in the gamma range in a sample of monolingual and bilingual speakers, and then used the computation model described in our earlier work (Tichko & Skoe, 2017) to simulate whether variation in subcortical source activation could influence the location of the GRmax (i.e., the frequency of the GRmax). To help define the GRmax, stimuli above the gamma range were also tested. Based on set theory, the maximum is, by definition, the largest in a set. Therefore, while the other stimulus frequencies in our stimulus set were not considered directly in the current analysis, they are necessary for defining the maximum.

2. Methods

2.1. Experimental Design and Subjects

Forty-four healthy young adults (mean age = 20.4; standard deviation = 1.3 years; 29 female) participated in the study. All procedures were approved by the University of Connecticut's Institutional Review Board and written informed consent was obtained in English for each participant. All participants were screened to confirm that they had no history of neurological disorders (e.g., seizures, cerebral palsy, spina bifida, or any syndrome associated with the central or peripheral nervous systems). Participants were also confirmed to have normal otoscopy, clinically normal bilateral air conduction thresholds <=20 dB HL for octaves from 125 to 8000 Hz, and click-evoked auditory brainstem response (ABR) wave latencies within normal limits for a 70-dB nHL rarefaction click presented at 31.25 Hz (using norms reported in Skoe et al., 2015). Collectively, this helped to confirm that participants had normal audiometric function.

In addition to language learning history, participants also reported on their musical training history, specifically their total years of music training and the age they started training. The first author initially approached the data analysis without knowledge of the participants' auditory experiences. To ensure objectivity, he remained unaware of the participants' bilingual or monolingual status during the initial analysis phase. It was only after establishing the stimulus frequency that yielded the maximum FFR, termed the GRmax, that the participants' group membership for subsequent analyses was revealed to the first author. This blind analysis was crucial for preventing any bias that might stem from knowledge of the participants' language experiences, thereby upholding the integrity of the data interpretation process.

2.2. Grouping Criteria

Participants were grouped into English monolinguals (*n* = 20, 12 females, mean age = 19.1; standard deviation = 1.4 years;) and bilinguals (*n* = 24, 16 females, mean age = 20.1; standard deviation = 1.2 years; 16 females) based on their self-reported language exposure. Participants were classified as "bilingual" if they self-reported learning English plus another language before age 15. Bilingual participants reported speaking English in combination with another language. The non-English languages represented in the sample are Arabic, Bangla, Bulgarian, Cantonese, Creole, Dutch, French, Gujarati, Hebrew, Hindi, Italian, Mandarin, Marathi, Russian, Serbian, Spanish, Tagalog, Tamil, Turkish, and Urdu.

Most of the bilingual participants (n = 21) learned the non-English language from their parents or family members and were "early" bilinguals (Kovelman et al., 2008) in that they started acquiring both languages before age 6. Three participants began learning their second language later in adolescence (13 years old, n = 2; 14 years old, n = 1). These three participants self-identified as bilinguals and were routinely in bilingual environments at the time of testing. Despite their later start, their exposure levels to a second language and proficiency were comparable to other bilinguals, indicating significant immersion and use of both languages.

In addition to the age of language acquisition (mean age = 3.84; standard deviation = 4.11 years), bilingual participants rated their ability level using a 10 point scale (1 = lowest, 10 = highest; mean points = 8.45; standard deviation = 0.46 points), and their exposure to each language in three-year intervals (i.e., from birth to current exposure) (Table 1).

2.3. FFR Stimuli and Recording Protocol

FFRs were recorded to 200-ms stimuli reflecting six unique frequencies in the low gamma range, plus seven additional frequencies. A total of 13 different F0s were tested (F0 = 29, 32, 43, 51, 61, 73, 87, 100, 110, 155, 207, 256, 293 Hz; rounded values). Each of the 13 F0s maps onto a different note on the piano. These frequencies were selected from the 91 used in our previous work to capture the stimulus frequencies where the FFR has local amplitude maxima and minima. Each stimulus was presented in a separate block with a presentation rate of 2.5/second using a protocol described in Tichko and Skoe, 2017. The average testing time was ∼ 1 h (4/min per stimulus). The stimuli were triangle waves that contained only the fundamental frequency (F0) and odd harmonics. For stimuli above 100 Hz, the stimulus intensity was calibrated to be 80 dB LAeq. For the lower stimuli, the calibrated intensity was 64, 64, 68, 69, 75, and 77 dB LAeq (rounded), for 29, 32, 43, 51, 61, 73, 87 Hz (respectively). The intensity was lowered in this range for greater listening comfort. Small differences in stimulus intensity are unlikely to influence our findings as all stimulus intensities were suprathreshold (Bidelman and Powers, 2018), and the intensity was fixed for a given stimulus.

We used the Psychophysics Toolbox extensions in MATLAB to deliver the stimuli. Stimuli were presented to the right ear using Mu-metal electromagnetically (EM) shielded tube phones (ER-3, Intelligent Hearing Systems, Inc.). Each stimulus was presented a total of 600 times to obtain at least 500 artifact-free responses. The stimulus was digitized by a USB-powered external sound card (M−Audio M−Track) at 44.1 kHz/16-bit resolution, and the audio stimulus was routed through a StimTrak box (Brain Products GmbH) into an auxiliary channel of the EEG amplifier (actiCHamp amplifier, Brain Products GmbH). This auxiliary channel provided accurate stimulus timing information that was used to average responses across trials.

FFRs were recorded using a 1-channel ipsilateral array with three Ag-AgCl electrodes (Multitrodes, Brain Products GmbH) placed on the scalp at Cz, FPz, and A2 (according to the 10–20 electrode system). Before electrode placement, the electrode sites were mildly exfoliated to produce impedances under 5 kOhms. The electrodes were plugged into a bipolar pre-amplifier with a gain factor of 50 (EP-PreAmp module, Brain Products GmbH). The actiCHamp amplifier interfaced with the bipolar pre-amplifier through a splitter box. BrainVision Recorder software (Brain Vision GmbH) recorded the electrical signals using a 25-kHz sampling rate and 0.05 mV/bit resolution (−409 to + 409 mV range) with no online filters applied. The FFR recording was made inside an electromagnetically shielded sound booth. To create a relaxing atmosphere, participants sat in a recliner chair and watched a self-selected movie that played silently with English captions. The movie was projected onto the booth wall from a ceiling-mounted LCD projector outside the booth window.

Table 1
Amount of second language (L2) exposure, in percent %, across the lifespan from birth to age 21. Mean and median values represented for all bilinguals. For each age range, the group minimum and maximum are also reported. There was as general trend for participants to have increased levels of L2 exposure throughout childhood into adolescence.

L2 Exposure (% of time) by Age							
Age	**0 – 3**	**3 – 6**	**6 – 9**	**9 – 12**	**12 – 15**	**15 – 18**	**18 – 21**
Mean	21 %	37 %	49 %	54 %	55 %	58 %	50 %
Median	10 %	40 %	52 %	50 %	50 %	70 %	30 %
Minimum	0 %	0 %	3 %	3 %	4 %	8 %	0 %
Maximum	56 %	81 %	100 %	100 %	100 %	100 %	100 %

Custom MATAB scripts were used to process the recordings. The neural responses were offline filtered from 15 to 10,000 Hz using a third-order, one-dimensional Butterworth filter. Notch filters were applied at 60 Hz, 1 kHz, 2 kHz, and 4 kHz using a second-order IIR notch filter design with a Q-factor of 35 to remove a small amount of environmental electrical noise that was present at these recording frequencies. (Note that these frequencies do not coincide with any stimulus frequencies). The auxiliary EEG channel provided the timing of the stimulus onset, which was used to correct the acoustic transmission delay introduced by the tube connecting the transducer and foam ear insert. For each stimulus, 500 artifact-free trials were averaged to create a single FFR waveform after discarding trials that exceeded ± 60 μV.

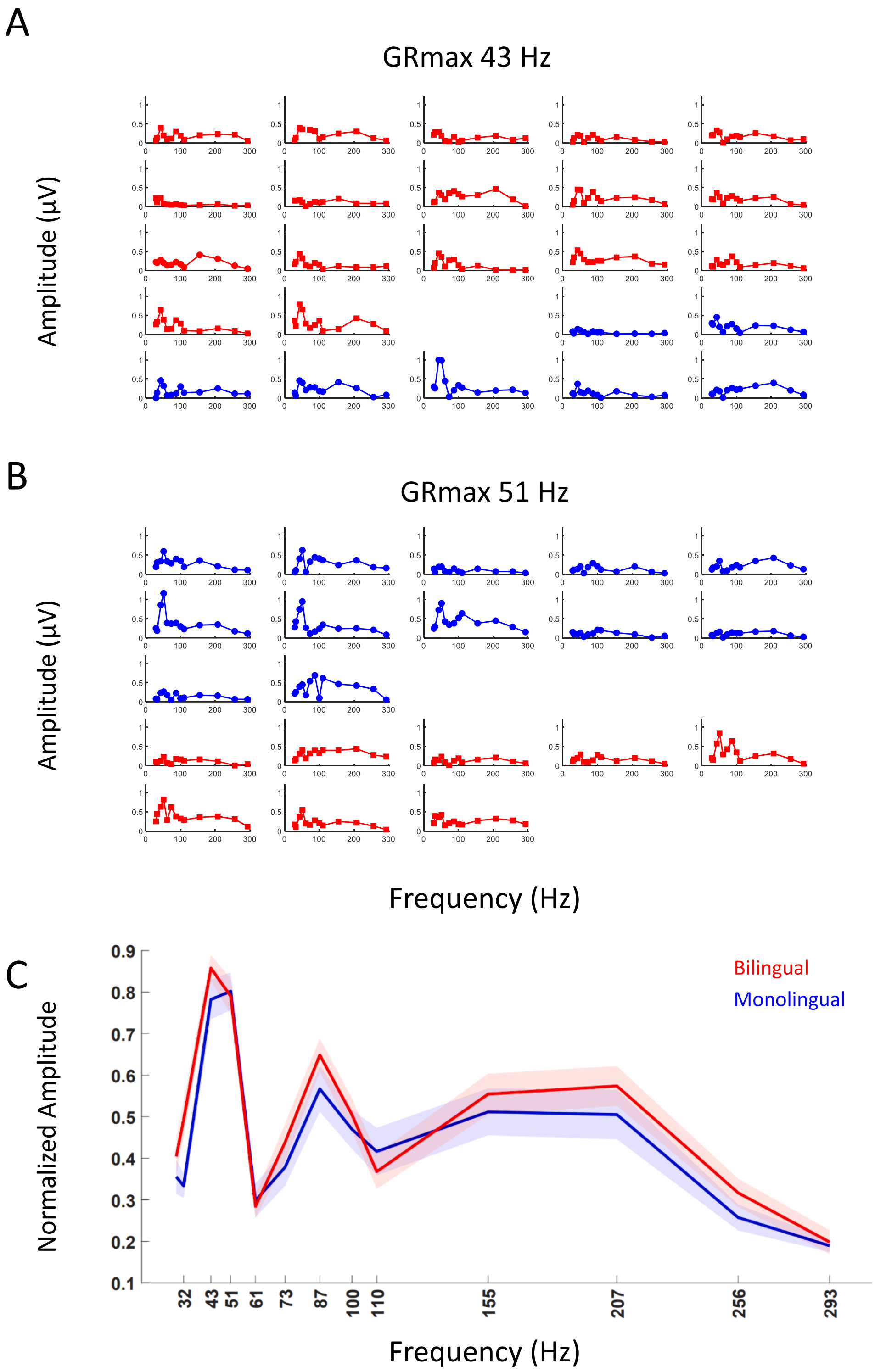

Fig. 1. Panels (A) and (B) show plots for individual participants, with color coding indicating whether the participant was bilingual (red) or monolingual (blue). Note that most of bilinguals had their GRmax at 43-Hz (A) and most monolinguals had their GRmax at 51-Hz (B). Panel (C) shows the group averages for bilinguals and monolinguals, using a normalized amplitude y-axis. The amplitudes were normalized for each participant before group averaging; the normalization process involved dividing the amplitude for each frequency by the amplitude of the GRmax frequency for eah participant. Shaded areas represent +/- 1 standard error of the mean.

The averaging window was 350-ms, which included 50 ms of pre-stimulus activity and 100 ms of activity after each stimulus.

A fast Fourier transform (FFT) was applied to each waveform. The FFT was performed on a 250-ms section of the response (0–250 ms) after applying a 250-ms Hanning window. The output was scaled to microvolts by taking the absolute value and multiplying the result by two divided by the signal length. Zero-padding to 1 s was applied as part of the FFT to increase the sample rate of the spectral estimate. From the spectral estimate of the FFR, the response amplitude at the F0 frequency of each stimulus was extracted. We then identified the location of the GRmax, i.e., the stimulus frequency that produced the strongest response in the gamma range. Our data-driven analyses focused specifically on the stimulus frequency that produced GRmax; other dimensions of the FFR were not considered. Exploratory analyses were performed using IBM SPSS Statistics (version 28).

3. Results

3.1. Gamma range variation: blinded dataset

Consistent with what has been seen in other studies (Coffey et al., 2019, Bidelman, 2018, Clinard et al., 2010), the magnitude of the FFR to

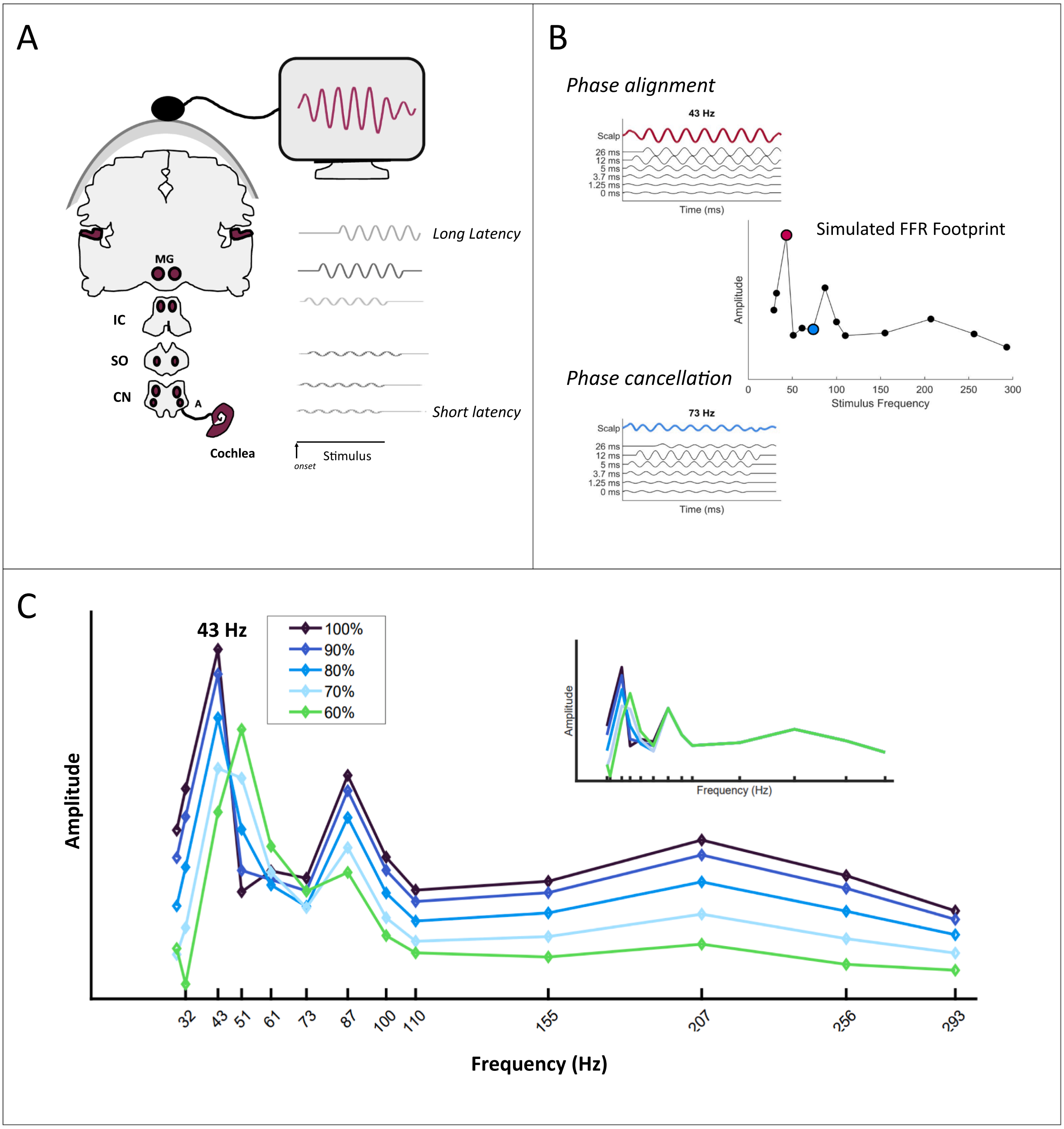

Fig. 2. (A) and (B): Model schematic to illstrate how six theoretical auditory generators along the auditory pathway (shown in panel A), with different starting latencies (0, 1.25, 3.7, 5, 12, 26 ms, respectively) and amplitudes, can mathematically combine (i.e., via linear superposition) to create composite frequency-following responses (FFRs) signals at the scalp that vary in amplitude as a function of stimulus frequency, creating the local maxima (e.g., 43 Hz in red) or minima (e.g., 73 Hz in blue) seen in panel B. (C) Simulated FFRs illustrating the impact of decreased activation of early latency generators (0, 1.25, 3.7 5 ms). Across the four lines, the activation drops from 100 % to 60 %, and with it, the maximum at 43 Hz shifts to 51 Hz, suggesting that a GRmax at 51 Hz reflects relatively less activity from peripheral and subcortical generators. The inset panel shows a similar set of simulations but where the drop in activation is limited to stimulus frequencies below 75 Hz to align the model with the empirical FFR data seen in Fig. 1.

the stimulus F0 varied systematically in all participants as a function of stimulus frequency. FFR local maxima emerged ~ 43, 51, 87, and 207 Hz, and minima emerged ~ 29, 61, 110, and 293 Hz (Fig. 1). Most interestingly, in the gamma range, individual variation was also observed; roughly half of the participants ($n = 24$, Fig. 1-**A**) showed a GRmax located at 43 Hz, with the other half ($n = 21$, Fig. 1-**B**) at 51 Hz. In those with the GRmax located at 43 Hz (Fig. 1-**A**), a dependent *t*-test confirmed that the response at 43 Hz ($M = 0.404$; $SD = 0.193$) is indeed larger in spectral amplitude than the response at 51 Hz ($M = 0.303$; SD = 0.196; $p \leq 0.0001$). In the same fashion (Fig. 1-**B**), in those with the GRmax located at 51 Hz ($M = 0.490$; $SD = 0.305$), the spectral amplitude was larger than the response at 43 Hz ($M = 0.359$; $SD = 0.233$; $p \leq 0.0001$).

3.2. GRmax and Bilingualism

After observing this individual variability in the GRmax location in the blinded dataset, we then explored whether there was a statistically significant relationship between the location of the GRmax (not the absolute amplitude) and bilingual language experience. A chi-square test for independence indicated a significant association between the location of the GRmax and language experience with a moderate effect size, $\chi 2$ (1, $n = 44$) = 4.227, $p = 0.04$, $phi = 0.31$. Next we applied a binary logistic regression model to study whether language experience affected the probability of whether the GRmax was located at 43 Hz vs. 51 Hz, after first performing an Omnibus test to confirm the appropriateness of this model for this dataset ($\chi 2 = 4.281$, $df = 1$, $p = 0.039$). The regression model included one dichotomous independent variable: language experience (either monolingual or bilingual). The dependent variable was the GRmax location (either 43 Hz or 51 Hz). Language experience emerged as a predictor for the GRmax location ($p = 0.043$). Among the 24 participants displaying the 43 Hz GRmax profile (Fig. 1-**A**), 70.8 % ($n = 17$) were bilingual, and of those with the 51 Hz profile, 60.0 % ($n = 12$) were monolingual (Fig. 1-**B**). In other words, bilingual participants were 2.7 times more likely to have their GRmax at 43 Hz, whereas monolingual individuals were more likely to have their GRmax at 51 Hz (Fig. 1-**A, B & C**).

The results of the Chi-Square and regression analysis prompted an exploration of the experiential variables that might drive whether a bilingual has their GRmax at 43 vs. 51 Hz. Specifically, we tested whether bilinguals with the GRmax at 43 Hz differed from the bilinguals with the GRmax at 51 Hz, with respect to the age that the second language was acquired, the self-rated ability to speak the first and second language, the amount (%) of exposure to a second language before age six; the degree of decline in first language exposure (comparing exposure before age six to present age); and also the amount of present-day exposure to both languages (i.e., how “bilingual” they currently are). In estimating the current level of bilingualism, we focused on the non-English language (NEL), irrespective of whether it was self-reported as the first or second language, given that our sample comprised college students residing and studying in an English-speaking country. The groups of bilinguals were compared using an independent Welch’s T-test to account for the unequal variance (Levene’s test for equal variance, $p = 0.006$). The results indicated no significant differences between the two bilingual groups with respect to the age of second language exposure or current ability. However, differences were noted for both current language exposure and change in language exposure over time. The bilinguals with the 51 Hz profile had, on average, a greater decline in exposure to the non-English language over time compared to the bilinguals with the 43 Hz profile: 54.8 % ($M = 54.8$, $SD = 23.4$; Hedge’s $g = 0.725$, CI 95 % −1.56 to 0.12) vs. 36.0 % ($M = 36.0$, $SD = 25.7$) (t (15.4) = 1.7, $p = 0.093$, CI 95 % [-41.11 to 3.49]). In addition to showing a greater decline, the bilinguals with the 51 Hz profile were currently “less” bilingual (Hedge’s $g = 1.05$, CI 95 % 0.16 to 1.92), in that they had less current exposure to the non-English language ($M = 8.6$ %, $SD = 9.54$) compared to the bilinguals with the 43 Hz profile ($M = 37.9$ %, $SD = 31.7$; $t(19.5) = 3.3$, $p = 0.03$; CI 95 % [11.27 to 47.35]).

We also evaluated musical history variables, specifically the total years of musical training and the age at onset of training. Despite observing that, on average, the 43 Hz GRmax group ($M = 4.7$, $SD = 4.0$) had engaged in musical training for more years (51 Hz; $M = 2.5$, $SD = 2.7$) and began their musical education ($M = 8.4$, $SD = 3.7$) earlier than the 51 Hz GRmax group ($M = 10.7$, $SD = 3.0$), statistical analyses revealed no significant differences between the two GRmax groups in either the total years of musical training (U = 306.000, $p = 0.063$) or the age at onset musical training $t(33) = 1.8$, $p = 0.075$.

3.3. Computational FFR model

Our analyses above suggest that the level of bilingualism may influence the location of the GRmax. For a possible neuroanatomical explanation of this finding, we employed a theoretical model of the FFR described in detail previously and summarized briefly here. For specifics, refer to Tichko and Skoe (2017). Similar modeling approaches were presented in Purcell et al., 2004 and Easwar et al., 2018. Our model, which builds from a previous model developed by Gardi et al., 1979, included six putative generators of the FFR, with first spike latencies chosen to align with physiological properties that approximate cochlear hair cells (that generate the cochlear microphonic), the cochlear nucleus (CN), superior olive (SOC), inferior colliculus (IC), primary auditory cortex (PAC), and non-PAC (0, 1.25, 3.7, 5, 12, 26 ms, respectively). In the model, the phase-locked output of each generator is modeled as a sine wave that matches the stimulus frequency and whose latency and amplitude are adjusted to be progressively higher from the first (most peripheral) to the sixth (most central) generator (Fig. 2). For the two generators with the longest latencies, the model included a drop out frequency, to represent the upper limit of their respective phase-locking ability. The drop out frequency was 200 Hz for the 12 ms generator (with a gradual tapering of the amplitude from 75-200 Hz), and 100 Hz for the 26 ms generator (with a gradual tapering of the amplitude from 50-100 Hz). The six sinewave outputs were then summed to generate a composite representing the volume-conducted FFR signal recorded at the scalp. For an illustration, see Fig. 2. By varying the amplitude of a single or a set of putative generators, it is possible to simulate how the scalp-recorded FFR might change when one or more generators are potentiated or attenuated.

Here, we were specifically interested in the degree to which the early latency (brainstem) generators might account for the location of the GRmax. To model the GRmax at 43 Hz, we used a base model where the latencies were set to 0, 1.25, 3.7, 5, 12, 26 ms and the relative contribution of each generator was 6 %, 6 %, 10 %, 18 %, 30 % and 30 %. As described in Tichko and Skoe (2017), this base model produces a local maximum at 43 Hz. Then, keeping the latencies of the base model fixed, we modeled the GRmax at 51 Hz, by selectively reducing the amplitude of the four earliest latency generators that align with subcortical onset activation times. The amplitude was reduced step-wise from the base model, in decrements of 10 % (Fig. 2-**C**). In the simulations we only adjusted the amplitude, as it allows us to observe the relative contribution of subcortical generators to the final aggregated response. Fig. 2-**C** shows how GRmax peak shifts in frequency as the amplitude decreases. Specifically, when these four generators each decrease their amplitude by 60 %, the GRmax peak shifts to 51 Hz. In the previous step (a 70 % decrease), the FFR footprint had two maxima in the gamma range with roughly equivalent amplitude at 43 and 51 Hz. To match the simulations more closely to the empirical FFR data in the speech range, we then constrained the amplitude reduction to the gamma-related frequencies (Fig. 2-**C, inset**).

4. Discussion

The “40-Hz phenomenon” has been extensively investigated through various lenses, including epidemiology, diagnostics, prevalent pathological models, and contemporary treatment targets (Galambos et al., 1981, Artieda et al., 2004, McDermott et al., 2018). Our study examines a specific dimension of this phenomenon, namely the stimulus frequency that produces the strongest (maximal) response in the vicinity of 40 Hz (30–60 Hz, gamma range). To focus on the phenomenon and not its various terminology, we refer to it as the GRmax. While prior studies have recognized that individuals vary in which stimulus produces the GRmax, none have, to our knowledge, considered the role of auditory experience in influencing this characteristic of the FFR. Our exploratory analysis suggests the possibility that enriched exposure to sound (bilingualism) could drive experience-dependent plasticity in subcortical (i.e., early latency) auditory regions to influence whether the scalp-recorded FFR is stronger at a lower vs. higher frequency in the gamma range.

In this dataset, we recorded FFRs to a total of 13 stimulus frequencies (29 Hz to 293 Hz), and found, as expected from the literature, that the gamma range produced particularly strong FFRs, but with notable individual variation in whether 43 or 51 Hz produced the strongest response, with a roughly 50–50 split across the dataset (Fig. 1). An interesting trend was observed: we found that individuals with bilingual backgrounds tended to exhibit their GRmax at 43 Hz, whereas monolingual individuals predominantly showed their GRmax at 51 Hz. Experience-dependent plasticity of the FFR for speech frequencies (>100 Hz) is well documented (e.g., Song et al., 2008, Krishnan et al., 2010, Zhao et al., 2018), but little attention has been given to lower stimulus frequencies, especially in investigations of bilingual populations. Previous studies using speech stimuli with F0s ~ 100 Hz found that bilinguals have stronger FFRs to the F0 of a speech stimulus (100 Hz) than monolinguals (Krizman et al., 2012, Skoe et al., 2017, Skoe et al., 2022) and that the extent of the enhancement correlated with their ability to inhibit irrelevant stimulation (Krizman et al., 2012) and the amount of bilingual experience (Krizman et al., 2015). These bilingual FFR enhancements have been attributed to heightened subcortical activation (relative to monolinguals), as subcortical auditory areas are the dominant, though not exclusive, contributor to FFRs to sounds in the 100 Hz range (Krizman et al., 2012, Krizman et al., 2015, Giroud et al., 2020). Based on this, we wondered whether differential activation of subcortical areas might be a plausible explanation for the GRmax being at a lower frequency for most of the bilinguals in our sample. To test this, we simulated the effect of altering the strength of subcortical activity using a computation model of FFR generation (Tichko and Skoe, 2017, Gardi et al., 1979). This model is based on one first proposed by Gardi et al., 1979 and was informed by data from subsequent studies, such as Kuwada et al., 2002.

Our simulations support a subcortical explanation of the group difference: when early latency sources are stronger, the GRmax of the model falls at 43 Hz, and when the early-latency sources are weaker, the GRmax shifts to 51 Hz. Interestingly, we find that bilinguals with greater current exposure to both languages are more likely to have their GRmax at 43 Hz, and those with less current exposure are more likely to have their GRmax at 51 Hz. Taken together with our model simulations, this finding supports the theory that bilingualism modulates the strength of the subcortical response to sound.

The simulations here are intended to be purely illustrative and purposefully reductive; they are not intended to be exhaustive (e.g., probe every model adjustment or capture complex aspects of basilar membrane nonlinearities; Dau et al., 2003), nor are they intended to be specific (e.g., simulate the FFR footprint of a specific individual in the dataset) or be definitive with respect to specific sources. Indeed, the latencies and amplitudes used in the model are relative and not absolute — they define the relative relationship among a set of possible generators that roughly approximate more peripheral to more central sources without making any definitive claims that there are only six generators. In theory, two weak sources could have such similar latencies (i.e., and be in close phase alignment) that they effectively act as one in generating the FFR at the scalp (assuming similar scalp topographies). For instance, the first spike latency of the auditory nerve overlaps closely in time with the cochlear microphonic. It could also be argued that the longer generators might include the thalamus-cortical pathway, as suggested by Lerud et al. 2023 and Coffey et al., 2019. Although the model is inherently simplified, it is also straightforward to interpret, with the capacity to provide new insight into complex phenomena that can be tested using more sophisticated (and expensive) neuroimaging source estimating techniques (e.g., high-density EEG together with individual structural scans) and refined using larger datasets.

In the last decade, there has been growing interest in the sources of the FFR (reviewed in Coffey et al., 2019, Gorina-Careta et al., 2021, Lerud et al., 2023), much of which has centered on whether there is significant cortical activation in the speech frequency range and/or whether EEG and MEG yield similar source estimates (Bidelman 2018). For example, using high-density EEG, our recent work (Lerud et al., 2023) showed evidence of both subcortical and cortical activation for both low (80 Hz) and high (210 Hz) stimulation frequencies, supporting the multiple generator hypothesis of the FFR. This hypothesis posits that both cortical and subcortical auditory areas contribute to the scalp-recorded response but with the precise blend and proportion of sources fluctuating based on the characteristics of the stimulus, the recording approach (e.g., EEG vs. MEG; EEG electrode configurations) and the individual being tested. Recent studies of the FFR (and ASSR) have overlooked what we believe is an important further contributing factor: the participant’s demographics (e.g., age) and auditory experiences (e.g., spoken language experience, musical training). That is, individual-level factors, we argue, may influence the specific proportion of sources.

Our findings highlight the importance of using multiple stimuli, considering the entire auditory system, and factoring in a person’s auditory experience when seeking a neurobiological explanation for individual differences in the FFR. Indeed, testing more finely spaced frequencies between 43 and 51 Hz and using higher-density electrode arrays is likely to provide an even more nuanced account of how activity from different sources blends at the scalp to create the FFR to different stimulus frequencies. Conventionally, cortical regions are considered the dominant FFR source for stimulus frequencies < 80 Hz, and the involvement of subcortical sources is often overlooked. Similarly, until recently, cortical involvement in higher frequencies was also underplayed. To illuminate the full biodiversity of the FFR, it may be important to consider the role of both cortical and subcortical sources, regardless of the stimulus frequency. Thus, differential responsiveness of the cortex could potentially (but does not necessarily) explain individual differences to speech (e.g., Hartmann et al., 2019). Likewise, differential brainstem responses could explain variation at lower frequencies. Auditory areas undergo significant age-related changes, especially in early life, with a theorized critical period for brainstem regions between birth and adolescence (Skoe and Kraus, 2013). Intensive exposure to two spoken languages both during and outside this time window could potentiate the response from subcortical structures (Liu et al., 2023), leading to the observed trends. Our findings from this exploratory study open the door to more systematic investigations on how early or later life language experience might individuate activity in the gamma range and above.

When studying bilingual populations, many variables can be considered (age of acquisition, mode of acquisition, ability, amount of exposure, and the specific languages learned, to name a few) (De Bruin, 2019). In the current study, we included a heterogeneous sample of bilinguals with a mixture of different languages and learning histories to capture a wide representation of bilingualism. Observing common patterns across bilinguals of such experiential breadth is noteworthy as it suggests a generalized phenomenon relating to sound enrichment. Seeing a similar effect for musicianship supports this possibility. In the current study, language ability was based on self-report, which may overestimate or compress variation in second language ability across the bilinguals. Focusing future work on a single language pair and a more limited range of language age of acquisition may also yield clearer divisions between monolinguals and bilinguals.

Our exploratory analysis suggests that bilinguals with less/decreased bilingual exposure are more likely to pattern like monolinguals. However, it is currently unclear why some monolinguals pattern like bilinguals. In these cases, other life experiences or biological processes not accounted for here could be at play. Also, while we focus on subcortical involvement as a possible explanation of differences between bilinguals and monolinguals (Krizman et al., 2012, Liu et al., 2023), this does not preclude the possibility of the FFR being influence by both subcortical and cortical plasticity. Indeed, other research suggests that bilinguals may also have stronger onset activation in auditory cortex to tonal stimuli (Vihla et al., 2002) and more consistent cortical responses to sound (Krizman et al., 2015, Krizman et al., 2014). Other works suggests that auditory cortex might be larger in bilinguals (Ressel et al. 2012), and that bilinguals may engage different auditory, linguistic, and cognitive networks than bilinguals (e.g., Kovelman et al., 2008; Jasinska et al., 2013). In the current investigation, the scope was intentionally limited to the gamma range and the GRmax; in future work, we will consider the higher frequencies in a larger dataset in which language ability, learning history, and auditory and cognitive ability are better controlled. Thus, currently the functional implications of our findings are unclear.

5. Conclusion

The auditory system, from single cells to sophisticated auditory networks, has been widely investigated in humans and experimental animals over the last three decades (Irvine, 2017). This work has revealed individual differences in the structure and function that relate to variations in auditory abilities, age, the acoustic demands and dynamics of the individuals' sonic environment, and the communication requirements and experiences of the individual (Kilgard, 2001, Gordon et al., 2003, Kraus & White-Schwoch, 2015, Skoe et al., 2017). Over this same period, phasic activity in the 40–50 Hz range has received significant attention. Yet, the potential role of language experience has not, to our knowledge, been investigated. Here, in this exploratory data-driven study, we provide evidence that language experience may alter the source profile of the FFR and contribute to the previously observed individual differences in phase-locked responses to low-frequency stimuli in the gamma range. This evidence, while preliminary and just below the conventional threshold for statistical significance, underscores the need for future research to corroborate the relationship between gamma activity and language experience. Collectively, our findings emphasize the importance of examining the role of language experience in shaping auditory processing and the value of using multiple stimuli. They also highlight that the FFR, by its nature, is a far-field response with multiple sources. Further studies with functional and structural imaging techniques are needed to unravel the precise mechanisms underlying these observed differences and their implications for language processing and communication.

CRediT authorship contribution statement

Luan Tonelli: Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, Validation. **Parker Tichko:** Methodology, Writing – original draft, Writing – review & editing. **Erika Skoe:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the study participants for volunteering their time. We thank Paige Kingsley, Hannah Waldo, Kelly Linehan, Erica Scarpati, and Elizabeth Gernert for their assistance with data collection.

Research funds – ES.

Study design – ES, PT.

Data Collection – ES.

Data Analysis – LT.

Manuscript – LT, ES.

Funding for the study was provided by research startup funds (ES) provided by the University of Connecticut and NSF grant 1941147 (ES).

Data availability

Data will be made available on request.

References

Anderson, L. A., Hesse, L. L., Pilati, N., Bakay, W. M., Alvaro, G., Large, C. H., McAlpine, D., Schaette, R., & Linden, J. F. (2018). Increased spontaneous firing rates in auditory midbrain following noise exposure are specifically abolished by a Kv3 channel modulator. *Hearing Research, 365*, 77–89.

Artieda, J., Valencia, M., Alegre, M., Olaziregi, O., Urrestarazu, E., & Iriarte, J. (2004). Potentials evoked by chirp-modulated tones: A new technique to evaluate oscillatory activity in the auditory pathway. *Clinical Neurophysiology, 115*(3), 699–709.

Baltus, A., & Herrmann, C. S. (2015). Auditory temporal resolution is linked to resonance frequency of the auditory cortex. *International Journal of Psychophysiology, 98*(1), 1–7.

Batra, R., Kuwada, S., & Maher, V. L. (1986). The frequency-following response to continuous tones in humans. *Hearing Research, 21*(2), 167–177.

Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *NeuroImage, 175*, 56–69.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118*(2), 245–257.

Bones, O., Hopkins, K., Krishnan, A., & Plack, C. J. (2014). Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia, 58*, 23–32.

Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264*(1–2), 48–55.

Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications, 7*(1), 11070.

Coffey, E. B., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E., Zatorre, R. J., & Kraus, N. (2019). Evolving perspectives on the sources of the frequency-following response. *Nature Communications, 10*(1), 5036.

Dau, T. (2003). The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *The Journal of the Acoustical Society of America, 113*(2), 936–950.

De Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Science, 9*(3), 33.

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review, 130*(4), 953.

Easwar, V., Banyard, A., Aiken, S., & Purcell, D. (2018). Phase delays between tone pairs reveal interactions in scalp-recorded envelope following responses. *Neuroscience Letters, 665*, 257–262.

Farahani, F. V., Karwowski, W., & Lighthall, N. R. (2019). Application of graph theory for identifying connectivity patterns in human brain networks: A systematic review. *Frontiers in Neuroscience, 13*, 585.

Galambos, R., Makeig, S., & Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences of the United States of America, 78*(4), 2643–2647.

Galbraith, G. C. (1994). Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 92*(4), 321–330.

Gardi, J., Merzenich, M., & McKean, C. (1979). Origins of the scalp-recorded frequency-following response in the cat. *Audiology, 18*(5), 353–380.

Giroud, N., Baum, S. R., Gilbert, A. C., Phillips, N. A., & Gracco, V. (2020). Earlier age of second language learning induces more robust speech encoding in the auditory brainstem in adults, independent of amount of language exposure during early childhood. *Brain and Language, 207*, Article 104815.

Gorina-Careta, N., Kurkela, J. L., Hämäläinen, J., Astikainen, P., & Escera, C. (2021). Neural generators of the frequency-following response elicited to stimuli of low and high frequency: A magnetoencephalographic (MEG) study. *NeuroImage, 231*, Article 117866.

Gordon, K. A., Papsin, B. C., & Harrison, R. V. (2003). Activity-dependent developmental plasticity of the auditory brain stem in children who use cochlear implants. *Ear and Hearing, 24*(6), 485–500.

Gransier, R., Carlyon, R. P., & Wouters, J. (2020). Electrophysiological assessment of temporal envelope processing in cochlear implant users. *Scientific Reports, 10*(1), 15406.

Gransier, R., Hofmann, M., van Wieringen, A., & Wouters, J. (2021). Stimulus-evoked phase-locked activity along the human auditory pathway strongly varies across individuals. *Scientific Reports, 11*(1), 143.

Griškova-Bulanova, I., Živanović, M., Voicikas, A., Pipinis, E., Jurkuvėnas, V., & Bjekić, J. (2022). Responses at individual gamma frequencies are related to the processing speed but not the inhibitory control. *Journal of Personalized Medicine, 13*(1), 26.

Hari, R., Hämäläinen, M., & Joutsiniemi, S. L. (1989). Neuromagnetic steady-state responses to auditory stimuli. *The Journal of the Acoustical Society of America, 86*(3), 1033–1039.

Hartmann, T., & Weisz, N. (2019). Auditory cortical generators of the frequency following response are modulated by intermodal attention. *NeuroImage, 203*, Article 116185.

Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., & Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography, 15*, 69–86.24.

Irvine, D. R. (2018). Plasticity in the auditory system. *Hearing Research, 362*, 61–73.

Jasinska, K. K., & Petitto, L. A. (2013). How age of bilingual exposure can change the neural systems for language in the developing brain: A functional near infrared spectroscopy investigation of syntactic processing in monolingual and bilingual children. *Developmental Cognitive Neuroscience, 6*, 87–101.

Kilgard, M. P., Pandya, P. K., Vazquez, J., Gehi, A., Schreiner, C. E., & Merzenich, M. M. (2001). Sensory input directs spatial and temporal plasticity in primary auditory cortex. *Journal of Neurophysiology, 86*(1), 326–338.

Kraus, N., & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: A cognitive–sensorimotor–reward framework. *Trends in Cognitive Sciences, 19*(11), 642–654.

Kraus, N., & Nicol, T. (2017). The power of sound for brain health. *Nature Human Behaviour, 1*(10), 700–702.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25*(1), 161–168.

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics, 23*(1), 81–95.

Krizman, J., & Kraus, N. (2019). Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hearing research, 382*, Article 107779.

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. *Proceedings of the National Academy of Sciences, 109*(20), 7877–7881.

Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language, 128*(1), 34–40.

Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience Letters, 585*, 48–53.

Kovelman, I., Baker, S. A., & Petitto, L. A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Biling. Lang. Cogn., 11*(2), 203–223.

Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., & D'Angelo, W. R. (2002). Sources of the scalp-recorded amplitude-modulation following response. *Journal of the American Academy of Audiology, 13*(04), 188–204.

Lerud, K. D., Hancock, R., & Skoe, E. (2023). A high-density EEG and structural MRI source analysis of the frequency following response to missing fundamental stimuli reveals subcortical and cortical activation to low and high frequency stimuli. *NeuroImage*, Article 120330.

Lee, Y. S., Lueders, H., Dinner, D. S., Lesser, R. P., Hahn, J., & Klem, G. (1984). Recording of auditory evoked potentials in man using chronic subdural electrodes. *Brain, 107*, 115–131.

Liu, X., Gao, Z., Liu, W., He, X., & Wang, N. (2023). AoA-L2 and Usage-L2 modulate the functional neuroplasticity of the subcortex. *Brain and Language, 245*, Article 105323.

Luke, R., De Vos, A., & Wouters, J. (2017). Source analysis of auditory steady-state responses in acoustic and electric hearing. *NeuroImage, 147*, 568–576.

Mäkelä, J. P., & Hari, R. (1987). Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalography and Clinical Neurophysiology, 66*(6), 539–546.

McDermott, B., Porter, E., Hughes, D., McGinley, B., Lang, M., O'Halloran, M., & Jones, M. (2018). Gamma band neural stimulation in humans and the promise of a new modality to prevent and treat Alzheimer's disease. *Journal of Alzheimer's Disease, 65*(2), 363–392.

Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and Clinical Neurophysiology, 35*(6), 665–667.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences, 104*(40), 15894–15898.

Nozaradan, S., Schönwiesner, M., Caron-Desrochers, L., & Lehmann, A. (2016). Enhanced brainstem and cortical encoding of sound during synchronized movement. *NeuroImage, 142*, 231–240.

Parciauskaite, V., Voicikas, A., Jurkuvenas, V., Tarailis, P., Kraulaidis, M., Pipinis, E., & Griskova-Bulanova, I. (2019). 40-Hz auditory steady-state responses and the complex information processing: An exploratory study in healthy young males. *PLoS One1, 14*(10), Article e0223127.

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *International Journal of Audiology, 42*(4), 177–219.

Poulsen, C., Picton, T. W., & Paus, T. (2007). Age-related changes in transient and oscillatory brain responses to auditory stimulation in healthy adults 19–45 years old. *Cerebral Cortex, 17*(6), 1454–1467.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R. J., & Näätänen, R. (2001). Memory traces for words as revealed by the mismatch negativity. *NeuroImage, 14*(3), 607–616.

Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America, 116*(6), 3581–3593.

Ribary, U., Ioannides, A. A., Singh, K. D., Hasson, R., Bolton, J. P., Lado, F., Mogilner, A., & Llinás, R. (1991). Magnetic field tomography of coherent thalamocortical 40-Hz oscillations in humans. *Proceedings of the National Academy of Sciences, 88*(24), 11037–11041.

Ressel, V., Pallier, C., Ventura-Campos, N., Díaz, B., Roessler, A., Ávila, C., & Sebastián-Gallés, N. (2012). An effect of bilingualism on the auditory cortex. *The Journal of Neuroscience, 32*(47), 16597–16601.

Rodriguez-Fornells, A., De Diego Balaguer, R., & Münte, T. F. (2006). Executive control in bilingual language processing. *Language Learning, 56*, 133–190.

Rossi, E., Newman, S., Kroll, J. F., & Diaz, M. T. (2018). Neural signatures of inhibitory control in bilingual spoken production. *Cortex, 108*, 50–66.

Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: A tutorial. *Ear and Hearing, 31*(3), 302.

Skoe, E., Krizman, J., Anderson, S., & Kraus, N. (2015). Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex, 25*(6), 1415–1426.

Skoe, E., Burakiewicz, E., Figueiredo, M., & Hardin, M. (2017). Basic neural processing of sound in adults is influenced by bilingual experience. *Neuroscience, 349*, 278–290.

Skoe, E., Garcia-Sierra, A., Ramírez-Esparza, N., & Jiang, S. (2022). Automatic sound encoding is sensitive to language familiarity: Evidence from English monolinguals and Spanish-English bilinguals. *Neuroscience Letters, 777*, Article 136582.

Skoe, E., & Kraus, N. (2013). Musical training heightens auditory brainstem function during sensitive periods in development. *Frontiers in Psychology, 4*, 622.

Snyder, R. L., & Schreiner, C. E. (1984). The auditory neurophonic: Basic properties. *Hearing Research, 15*(3), 261–280.

Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience, 20*(10), 1892–1902.

Stapells, D. R., Linden, D., Suffield, J. B., Hamel, G., & Picton, T. W. (1984). Human auditory steady state potentials. *Ear and Hearing, 5*, 105–113.

Tichko, P., & Skoe, E. (2017). Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. *Hearing Research, 348*, 1–15.

Tsuzuku, T. (1993). 40-Hz steady state response in awake cats after bilateral chronic lesions in auditory cortices or inferior colliculi. *Auris, Nasus, Larynx, 20*(4), 263–274.

Uhlhaas, P. J., & Singer, W. (2012). Neuronal dynamics and neuropsychiatric disorders: Toward a translational paradigm for dysfunctional large-scale networks. *Neuron, 75*(6), 963–980.

Valle Rodrigues, M. V., Donadon, C., Guedes-Weber, M., Giorgi Sant'anna, S., Skarzynski, P. H., Hatzopoulos, S., Colella-Santos, M. F., & Sanfins, M. D. (2019). Frequency Following Response and Musical Experience: A Review. *Journal of Hearing Science, 9*(2).

Vihla, M., Kiviniemi, K., & Salmelin, R. (2002). Auditory cortical activation in Finnish and Swedish speaking Finns: A magnetoencephalographic study. *Neuroscience Letters, 322*(3), 141–144.

White-Schwoch, T., Nicol, T., Warrier, C. M., Abrams, D. A., & Kraus, N. (2017). Individual differences in human auditory processing: Insights from single-trial auditory midbrain activity in an animal model. *Cerebral Cortex, 27*(11), 5095–5115.

Zaehle, T., Lenz, D., Ohl, F. W., & Herrmann, C. S. (2010). Resonance phenomena in the human auditory cortex: Individual resonance frequencies of the cerebral cortex determine electrophysiological responses. *Experimental Brain Research, 203*, 629–635.

Zhao, T. C., & Kuhl, P. K. (2018). Linguistic effect on speech perception observed at the brainstem. *Proceedings of the National Academy of Sciences, 115*(35), 8716–8721.