

基于样本熵的听觉神经锁相机理的实验分析

朱莉 邓娟 吴建华 周南润

Experimental analysis of auditory mechanism of neural phase-locking based on sample entropy

Zhu Li Deng Juan Wu Jian-Hua Zhou Nan-Run

引用信息 Citation: [Acta Physica Sinica](#), 64, 184302 (2015) DOI: 10.7498/aps.64.184302

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.184302>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I18>

---

您可能感兴趣的其他文章

Articles you may be interested in

[复杂结构冲击声合成及实验验证](#)

[Impact sound synthesis method of complex structures and the experimental verification](#)

[物理学报](#).2014, 63(22): 224303 <http://dx.doi.org/10.7498/aps.63.224303>

[包络调制率和载波频率对听觉时间调制检测能力的影响](#)

[Effects of envelope modulation rate and carrier frequency on auditory temporal modulation detection](#)

[物理学报](#).2012, 61(16): 164302 <http://dx.doi.org/10.7498/aps.61.164302>

# 基于样本熵的听觉神经锁相机理的实验分析\*

朱莉<sup>1)</sup> 邓娟<sup>2)</sup> 吴建华<sup>1)</sup> 周南润<sup>1)</sup>†

1)(南昌大学信息工程学院, 南昌 330031)

2)(中国医学科学院, 北京协和医科大学生物医学工程研究所, 天津 300192)

(2015年3月12日收到; 2015年4月23日收到修改稿)

锁相是指系统的响应与周期性刺激的特定相位同步的物理现象。听觉神经的锁相揭示人的听觉认知基本的神经机理及改善听觉感知有重要意义。然而, 现有研究主要集中于心理物理方法和幅度谱分析, 不能有效区分包络响应和时域细节结构响应, 不能直观反映神经锁相。本文主要利用拔靴法和离散傅里叶变换, 提出了基于样本熵的时域细节结构频率跟随响应 (temporal-fine-structure-related frequency following response, FFR<sub>T</sub>) 的神经锁相值 (phase locking value, PLV) 计算方法, 用于分析神经物理实验数据。两个脑电实验结果表明: FFR<sub>T</sub> 的 PLV 样本熵显著大于包络相关频率跟随响应 (envelope-related frequency following response, FFR<sub>E</sub>) 的 PLV, 且二者正交独立, 新方法能有效地分别反映听觉系统对包络和时间细节结构的锁相机理; 基频处的响应主要来源于 FFR<sub>E</sub> 的锁相; 基频处, 不可分辨谐波成分包络的锁相能力优于对可分辨谐波; 基频缺失时, 畸变产物是不同的听觉神经通路的 FFR<sub>E</sub> 的混合; 谐波处, FFR<sub>E</sub> 集中于低频, FFR<sub>T</sub> 则集中于中、高频; 听觉神经元锁相能力与声源的频率可分辨性相关。FFR<sub>T</sub> 的 PLV 方法克服了现有 FFR 分析的局限性, 可用于深入研究听觉神经机理。

**关键词:** 听觉神经机理, 锁相值, 样本熵, 频率跟随响应

**PACS:** 43.64.-q, 43.60.Np, 89.70.Cf

**DOI:** 10.7498/aps.64.184302

## 1 引言

听觉是人脑最重要的认知之一, 也是人日常生活、学习和交流的主要方式<sup>[1]</sup>。人的听觉神经是一个复杂的上行与下行通路并行的系统, 揭示听觉认知神经机理广泛的基础研究和应用价值<sup>[1,2]</sup>。例如, 应用听觉感知和认知的特征可实现声源目标的自动识别<sup>[3]</sup>。人脑对音高、听觉空间等维度的感知可用于声源目标分类<sup>[4]</sup>。听觉频率选择机理可建立人脑的学习模型等<sup>[5]</sup>。由此可见, 听觉神经是计算神经科学、神经物理学、信息学及临床医学等多学科的研究热点<sup>[6,7]</sup>。人脑是一个非线性动力系统。近几年, 熵、混沌及随机共振等非线性理论已被应用于研究神经元放电活动中的非线性现象, 并对研究

神经机理起到了重要作用<sup>[8]</sup>。譬如, 利用等概率符号化样本熵区分注意力集中和发散两种状态的脑电信号<sup>[9]</sup>; 用相对转移熵研究正常脑电与癫痫脑电的不可逆性<sup>[10]</sup>; 应用非线性 Lyapunov 方法估计静态离散分布神经网络的状态<sup>[11]</sup>; 依据神经集群模型控制大脑的非线性特征, 并通过闭环控制抑制癫痫患者的神经活动<sup>[12]</sup>; 利用 Lyapunov-Krasovskii 函数计算延迟的稳态条件, 继而研究恒定延迟神经系统的渐进稳态问题等<sup>[13]</sup>。神经活动的锁相值 (phase locking value, PLV) 是指系统响应与周期性刺激的特定相位同步的物理度量。相位会影响神经信号的同步、相干等动力学特性。例如, 听觉神经的锁相揭示了听觉神经元选频特性; 相位噪声对 Hopf 分岔和鞍-结分岔点的神经兴奋性有不同的作用<sup>[14]</sup>。

\* 国家自然科学基金(批准号: 61463035, 61262084)、江西省自然科学基金(批准号: 20142BAB217022, 20122BAB211020)和江西省教育厅科技项目(批准号: GJJ14193)资助的课题。

† 通信作者。E-mail: znr21@163.com

声音可分为包络(envelope, ENV)和时间细节结构(temporal fine structure, TFS). 理论上, ENV反映信号慢变成分, 则样本熵小; TFS反映信号快变成本, 则样本熵大. 近年来, 听觉系统对ENV与TFS编码的生物物理特性及神经机理研究主要包括音高、噪声、掩蔽等. 当基频缺失时, ENV和TFS对音高感知的作用是听觉神经研究的热点<sup>[15]</sup>. 噪声会严重影响听力损伤患者对声音的TFS成分的识别, 而不影响对ENV的识别<sup>[16]</sup>. TFS语音识别率与噪声环境下捕捉原始语音的时间特征的能力呈线性相关. 在噪声环境下, 语音的TFS有助于研究听力损伤, 并提高助听器及人工耳蜗的助听效果<sup>[17]</sup>. 人在不同的刺激声下, 感知到的言语清晰度和音速主要取决于TFS而非ENV<sup>[18]</sup>. 以上研究主要用Gammatone滤波器重组等心理物理声学方法不能得到正交ENV和TFS分量, 分析其神经机理<sup>[19]</sup>. 以脑电(electroencephalography, EEG)信号分析为主的生理物理方法能更直观、量化地反映人脑的神经机理. 脑皮层下的听觉神经系统能够精确地复现声音刺激的基频、共振峰等特性, 可通过EEG记录下来, 即听觉脑干响应(auditory brainstem response, ABR). ABR可分为暂态-起始响应和稳态-频率跟随响应(frequency following response, FFR). 稳态FFR是听觉神经元对周期性刺激声源的锁相响应的集合. 现有的FFR研究主要集中于FFR的ENV幅度谱分析, 无法用于分析TFS锁相机理. 与此同时, 幅度谱分析受与频率呈反比的 $1/f$ 噪声以及工频干扰影响. 研究表明, 噪声环境中人的听觉认知主要取决于人脑对TFS的编码; 仅使用ENV信息人们的言语识别率会严重下降; 在人工耳蜗等听力辅助设备中增加TFS有助于提高听力障碍人士的言语识别率<sup>[20,21]</sup>. 由此可见, 研究TFS的锁相性质用重要的理论和临床应用意义.

本文作者在文献<sup>[22]</sup>中提出了计算包络相关FFR(envelope-related frequency following response, FFR<sub>E</sub>)的幅度谱信噪比(SNR)方法和锁相值(PLV<sub>E</sub>)方法. FFR<sub>E</sub>响应可通过计算刺激声源的基频及其谐波的幅度谱和PLV分析. 实验结果表明: 1)PLV<sub>E</sub>可直接解释FFR<sub>E</sub>的物理意义, 而解释幅度谱的物理意义及神经机理则需要估计特定频率下的测量噪声; 2)PLV<sub>E</sub>比幅度谱SNR的计算量小, 抗噪性好, 有效响应显著增加; 3)幅度谱SNR和PLV<sub>E</sub>可用于分析基频、可辨别谐波

及不可分辨谐波激励外周听觉通道的生理物理机理; 4)FFR<sub>E</sub>响应集中于基频和低频, 且基频处的FFR<sub>E</sub>受外界噪声影响更小.

为深入分析听觉神经锁相能力, 本文在文献<sup>[22]</sup>的基础上, 提出基于样本熵的时域细节结构频率跟随响应神经元锁相值(phase locking value of envelope-related frequency following response, PLV<sub>T</sub>)计算方法. PLV<sub>T</sub>与PLV<sub>E</sub>方法不同的是: 1)PLV<sub>T</sub>反映听觉神经对快变TFS的锁相, 而PLV<sub>E</sub>则反映其对慢变ENV的锁相; 2)PLV<sub>T</sub>的样本熵大, 而PLV<sub>E</sub>的样本熵小; 3)PLV<sub>T</sub>与PLV<sub>E</sub>二者样本熵应正交独立; 4)PLV<sub>T</sub>计算过程中正、负极性用反相运算, 而PLV<sub>E</sub>计算过程中正、负极性用同相运算. 本文在安静和噪声环境下, 设计了多组复合音刺激声源, 实现了听觉认知脑电实验; 计算不同刺激、不同环境下受试者实验数据的PLV<sub>E</sub>与PLV<sub>T</sub>, 比较二者的样本熵量值及正交独立性, 验证了新方法能够分别反映听觉神经系统对ENV和TFS的锁相能力; 并分析了在基频、基频缺失、可分辨谐波及不可分辨谐波等不同情况下FFR<sub>E</sub>和FFR<sub>T</sub>的锁相能力的统计特性. 本文通过分析ENV和TFS两种不同分量的FFR, 较全面地研究了不同刺激及不同环境下听觉神经通路放电锁相的神经机理.

## 2 基于样本熵的听觉神经锁相分析方法

### 2.1 数据预处理方法

首先, 将原始EEG实验数据经过80—3000 Hz的带通滤波(零相位, Hamming窗, 800阶次, FIR滤波器), 保留脑皮层下响应, 同时尽量减少低频噪声和脑皮层活动的干扰. 其次, 将滤波后信号依据刺激和响应时间分段. 若分段脑电中的眼动活动电位(electrooculogram, EOG)大于阈值, 则删除该分段脑电以消除眼动伪迹. 否则, 用左、右耳垂测量到的平均值作为参考信号, Cz处的响应与参考信号的差值则为真实响应. 本文中的实验, 以刺激声播放的起始时刻为时间零点, 每个分段脑电取值范围是从[-250 ms, -50 ms], EOG阈值设为60  $\mu$ V. 实验1的四种刺激条件下(正/负极性  $\times$  安静/噪声环境), 每个受试者无伪迹的有效分段脑电个数 $L \geq 1800$ ; 实验2中, 两种刺激条件下(正/负

极性  $\times$  安静环境), 每个受试者无伪迹的有效分段脑电个数  $L \geq 900$ . 为保证信号是稳态响应, 时域信号加 170 ms 时长的窗 (Slepian 窗函数, 4 Hz 频率分辨率, 最小谱泄漏), 时间起始点为刺激声音开始后的 10 ms.

## 2.2 听觉神经放电锁相值计算方法

### 2.2.1 计算 $\text{FFR}_T$ 锁相值 PLV

1) 针对每位受试者  $s$ , 每种环境  $x$ , 每种极性, 用拔靴法随机抽取 400 个试验作为样本, 形成样本集.

2) 计算每个抽样试验的时域响应 FFR 的离散傅里叶变换 (DFT). 受试者  $s$ , 第  $n$  个样本, 环境  $x$  下 FFR 的 DFT 记为  $S_{n,x}^s(f)$ , 其相位记为  $\phi_{n,x}^s(f)$ .

$$\phi_{n,x}^s(f) = \angle S_{n,x}^s(f), \quad (1)$$

其中,  $f$  表示频率.

3) 受试者  $s$ , 第  $i$  次抽样, 环境  $x$ , 计算  $\text{FFR}_T$  的 PLV, 记为  $\text{PLV}_{T,x,i}^s(f)$ .

$$\begin{aligned} & \text{PLV}_{T,x,i}^s(f) \\ &= \frac{1}{400} \left| \sum_{n \in \text{POS}_i} e^{j\phi_{n,x}^s(f)} - \sum_{m \in \text{NEG}_i} e^{j\phi_{m,x}^s(f)} \right|, \end{aligned} \quad (2)$$

其中,  $n \in \text{POS}_i$  表示第  $n$  个样本属于第  $i$  次抽样中的正极性样本集  $\text{POS}_i$ ,  $m \in \text{NEG}_i$  表示第  $m$  个样本属于第  $i$  次抽样中的负极性样本集  $\text{NEG}_i$ .

4) 每位受试者重复步骤 1)–3) 共 100 次, 生成  $\text{FFR}_T$  的 PLV 分布.

5) 计算每位受试者  $s$ , 每种环境  $x$  下  $\text{FFR}_T$  的 PLV, 记为  $\text{PLV}_{T,x}^s(f)$ .

$$\text{PLV}_{T,x}^s(f) = \frac{1}{100} \sum_{i=1}^{100} \text{PLV}_{T,x,i}^s(f). \quad (3)$$

### 2.2.2 计算 $\text{FFR}_T$ 的 PLV 显著有效人数

计算全部受试者的 PLV 均值, 并计算每个谐波点 PLV 显著大于随机数的受试者人数. 步骤如下.

1) 对所有条件、受试者和频率而言, 零假设模型是相同的, 生成惟一的 PLV 随机分布, 即  $\{\hat{R}^j\}_{1 \leq j \leq 1000}$ , 其中  $j$  表示第  $j$  次抽样. 计算  $\{\hat{R}^j\}_{1 \leq j \leq 1000}$  的相位是  $[0, 2\pi]$  范围内服从均匀分布的随机相位.

2) 重复 PLV 的步骤 1)–3) 共 1000 次.

3) 用  $\{\hat{R}^j\}_{1 \leq j \leq 1000}$  中大于或等于  $\text{PLV}_{T,x}^s$  的样本数占总样本数的百分比表示 PLV 可能出现的概率.

4) 根据实验的变量个数  $K$ , 设置阈值, 用 Bonferroni 方法修正. 若 PLV 的概率  $p < 0.05/K$ , 则认为 PLV 数值是显著有效的.

## 2.3 神经放电样本熵

为验证  $\text{FFR}_E$  和  $\text{FFR}_T$  的 PLV 能否独立反映大脑皮层下的听觉神经元核团的放电特性, 我们用样本熵作为评价指标. 样本熵是一种不计数的自身匹配的统计方法. 样本熵越大, 表示数列的复杂度越高, 信号快变; 相反, 样本熵越小, 表示数列的自身之间的相似程度越高, 信号慢变. 样本熵可以用少量的短数列估计出稳健的值, 抗干扰能力强, 适合生物信号的分析. FFR 信号持续时间短 (如本文的实验中, FFR 响应小于 200 ms), 数据量小, 适合用样本熵计算  $\text{FFR}_{\text{ENV}}$  和  $\text{FFR}_{\text{TFS}}$  的复杂性或称自相似性. 令每一位受试者 FFR 分量 ( $\text{FFR}_{\text{ENV}}$  和  $\text{FFR}_{\text{TFS}}$ ) 为一个数列  $g(t)$ ,  $t$  为时间点. 设维数为  $d$ , 相似容限为  $r$ , 则样本熵的计算步骤如下.

1) 提取时域 FFR 信号. 令潜伏期延时 10 ms, FFR 响应时长 170 ms, 响应时间起始点与每一个试验的刺激声音起始点对齐. 提取全部有用的分段脑电 (EOG 活动电位小于 60  $\mu\text{V}$ ) 的 10—180 ms 时间段内的响应. 第  $l$  个试验的正、负极性时域响应分别记为  $\text{POS}(l_{\text{POS}}, t)$ ,  $\text{NEG}(l_{\text{NEG}}, t)$ . 其中  $l_{\text{POS}} = 1, \dots, L_{\text{POS}}$ ,  $l_{\text{NEG}} = 1, \dots, L_{\text{NEG}}$ ,  $L_{\text{POS}}$ ,  $L_{\text{NEG}}$  分别为有用正、负极性分段脑电总数;  $1 \leq t \leq T$ ,  $T$  表示总时长.

2) 计算包络和 TFS 的时域平均 FFR, 分别记为  $g_E(t)$  和  $g_T(t)$ .

$$\begin{aligned} g_E(t) &= \frac{1}{2} \left( \sum_{l_{\text{POS}}=1}^{L_{\text{POS}}} \text{POS}(l_{\text{POS}}, t) \right. \\ &\quad \left. + \sum_{l_{\text{NEG}}=1}^{L_{\text{NEG}}} \text{NEG}(l_{\text{NEG}}, t) \right) \\ t &= 1, \dots, T, \end{aligned} \quad (4)$$

$$\begin{aligned} g_T(t) &= \sum_{l_{\text{POS}}=1}^{L_{\text{POS}}} \text{POS}(l_{\text{POS}}, t) \\ &\quad - \sum_{l_{\text{NEG}}=1}^{L_{\text{NEG}}} \text{NEG}(l_{\text{NEG}}, t) \\ t &= 1, \dots, T. \end{aligned} \quad (5)$$

3) 确定相似容限  $r$ . 本文按照脑电分析的常用规则, 令  $r$  为  $g_E(t)$  和  $g_T(t)$  两组时域响应的均方误

差最小值的0.2倍。

4) 按时间序列号分成 $d$ 维矢量, 即 $\{X_{dE}(1), X_{dE}(2), \dots, X_{dE}(T-d+1)\}$ 和 $\{X_{dT}(1), X_{dT}(2), \dots, X_{dT}(T-d+1)\}$ .

$$\begin{aligned} X_{dE}(t) &= [g_E(t), g_E(t+1), \dots, g_E(t+d-1)], \\ 1 \leq t &\leq T-d+1, \end{aligned} \quad (6)$$

$$\begin{aligned} X_{dT}(t) &= [g_T(t), g_T(t+1), \dots, g_T(t+d-1)], \\ 1 \leq t &\leq T-d+1, \end{aligned} \quad (7)$$

$X_{dE}(t)$ 和 $X_{dT}(t)$ 分别表示时间 $t$ 开始连续 $d+1$ 个时域平均FFR<sub>E</sub>和FFR<sub>T</sub>.

5) 定义第 $t_1$ 个和第 $t_2$ 个时间点的响应距离如下:

$$\begin{aligned} D[X_{dE}(t_1), X_{dE}(t_2)] \\ = \max[g_E(t_1+v) - g_E(t_2+v)], \end{aligned} \quad (8)$$

$$\begin{aligned} D[X_{dT}(t_1), X_{dT}(t_2)] \\ = \max[g_T(t_1+v) - g_T(t_2+v)], \end{aligned} \quad (9)$$

其中 $1 \leq t_1, t_2 \leq T-d+1; 0 \leq v \leq d-1$ .

6) 计算给定相似容限 $r$ 的模板匹配数. 所有 $1 \leq t \leq T-d+1$ 时间范围内, 计算距离 $D[X_{dE}(t_1), X_{dE}(t_2)], D[X_{dT}(t_1), X_{dT}(t_2)]$ 小于 $r$ 的数目, 令与总数 $T-d+1$ 的比值为 $B_{tE}^d(r)$ 和 $B_{tT}^d(r)$ .

$$\begin{aligned} B_{tE}^d(r) &= \frac{1}{T-d+1} \sum_u \left\{ \begin{array}{l} u = 1, \\ \text{if } D[X_{dE}(t_1), X_{dE}(t_2)] < r; \\ u = 0, \text{ otherwise} \end{array} \right\} \\ 1 \leq t_2 &\leq T-d, \quad t_2 \neq t_1, \end{aligned} \quad (10)$$

$$\begin{aligned} B_{tT}^d(r) &= \frac{1}{T-d+1} \sum_u \left\{ \begin{array}{l} u = 1, \\ \text{if } D[X_{dT}(t_1), X_{dT}(t_2)] < r; \\ u = 0, \text{ otherwise} \end{array} \right\} \\ 1 \leq t_2 &\leq T-d, \quad t_2 \neq t_1. \end{aligned} \quad (11)$$

再求 $B_{tE}^d(r)$ 和 $B_{tT}^d(r)$ 的时间 $t$ 平均值 $B_E^d(r)$ 和 $B_T^d(r)$ .

$$B_E^d(r) = \frac{1}{T-d} \sum_{t=1}^{T-d} B_{tE}^d(r), \quad (12)$$

$$B_T^d(r) = \frac{1}{T-d} \sum_{t=1}^{T-d} B_{tT}^d(r). \quad (13)$$

然后, 求更高维 $d+1$ 维矢量的给定相似容限 $r$ 的模板匹配数 $A_{tE}^d(r), A_{tT}^d(r)$ , 及其时间平均值 $A_E^d(r)$ 和 $A_T^d(r)$ :

和 $A_T^d(r)$ :

$$A_{tE}^d(r) = \frac{\sum_u \left\{ \begin{array}{l} u = 1, \text{ if } D[X_{(d+1)E}(t_1), X_{(d+1)E}(t_2)] < r; \\ u = 0, \text{ otherwise} \end{array} \right\}}{T-d-1} \\ 1 \leq t_2 \leq T-d, \quad t_2 \neq t_1, \quad (14)$$

$$A_{tT}^d(r) = \frac{\sum_u \left\{ \begin{array}{l} u = 1, \text{ if } D[X_{(d+1)T}(t_1), X_{(d+1)T}(t_2)] < r; \\ u = 0, \text{ otherwise} \end{array} \right\}}{T-d-1} \\ 1 \leq t_2 \leq T-d, \quad t_2 \neq t_1, \quad (15)$$

$$A_E^d(r) = \frac{1}{T-d} \sum_{t=1}^{T-d} A_{tE}^d(r), \quad (16)$$

$$A_T^d(r) = \frac{1}{T-d} \sum_{t=1}^{T-d} A_{tT}^d(r). \quad (17)$$

7) 计算 $d$ 点时间序列 $B_E^d(r), B_T^d(r)$ 与 $d+1$ 点时间序列 $A_E^d(r), A_T^d(r)$ 的相似概率, 即样本熵 $H_E(d, r)$ 和 $H_T(d, r)$ .

$$H_E(d, r) = \lim_{T \rightarrow \infty} \{-\ln[A_E^d(r)/B_E^d(r)]\}, \quad (18)$$

$$H_T(d, r) = \lim_{T \rightarrow \infty} \{-\ln[A_T^d(r)/B_T^d(r)]\}. \quad (19)$$

序列长度有限, 样本熵的计算允许用有限长的序列近似无限长序列, 因此可以用有限点样本熵估计值来表示时间序列的相似性.

### 3 FFR听觉认知实验

为验证FFR<sub>T</sub>幅度谱SNR和锁相值PLV分析方法的有效性, 并研究安静和噪声环境中听觉脑电信号的特性, 我们完成了两个听觉认知FFR实验, 实验均在美国Boston University的Auditory Neuroscience Laboratory完成.

#### 3.1 受试者

实验1包括16名受试者, 年龄为21—38岁, 其中4名女性. 实验2包括20名受试者, 年龄为20—30岁, 其中10名女性. 所有的受试者在实验前都进行了双耳纯音听力阈值测试, 测试频率为250—8000 Hz倍频程. 测试结果为, 在各个频率下所有受试者的单侧耳的听力阈值都低于15 dB, 且没有中枢神经或外周听觉损伤病史, 所有受试者听

力正常。受试者签署了同意实验程序协议，并给受试者支付了相应实验费用。

### 3.2 实验刺激声源

实验所用的刺激声音由 MATLAB 离线生成并存储，在实验中以 25 kHz 的采样率回放。实验 1 的刺激声源是复合音，由 8 个等强度的纯音的 sine 相位纯音 (100, …, 800 Hz) 组成，基频为 100 Hz。安静环境中，只播放此刺激声源。噪声环境中，加入了一个持续的高斯噪声 (经过 3000 Hz 低通)，信噪比 SNR 为 +10 dB。实验 2 包括 4 个不同的复合音，基频均为 100 Hz，每个复合音由等强度的 cosine 相位纯音组成，但是包括不同的谐波。第一个复合音包含 100 Hz 的前 5 个谐波 (低频: 100, …, 500 Hz)，对正常听力受试者而言，听觉外周神经很容易分辨全部五个谐波。第二个复合音包含 6—10 谐波 (中频: 600, …, 1000 Hz)，听觉外周神经能分辨这五个谐波中的一部分。第三个复合音包括 12—16 谐波 (高频: 1200, …, 1600 Hz)，听觉外周神经系统对这五个高次谐波的分辨率较差。第四个复合音包括全部前 20 个谐波 (宽频: 100, …, 2000 Hz)。

实验 1 和实验 2 的复合音长度均为 170 ms，包括 10 ms  $\cos^2$  窗函数的升降时间。声源刺激保持 80 dB 声强水平，同时向双耳播放。重复声源刺激的平均间隔为 770 ( $\pm 100$ ) ms，单个试验之间的时间扰动目的是为了保证与刺激无关的信号是随机相位。声音刺激以交错极性的方式播放，即一半试验是正极性，而另一半试验是负极性。

### 3.3 实验设备与电极放置方法

硬件设备：1)PC 机，通过编程用以控制实验中触发声音传输、存储数据等；2)声卡控制硬件 (System 3 Real-time Signal Processing Systems, Tucker Davis Technologies, Gainesville, FL)，控制声音传输过程；3)入耳式耳塞 (ER-1, Etymotic, Elk Grove Village, IL)。

记录软件是 BioSemi Active Two System (BioSemi, Amsterdam, Netherlands)，采样率为 16.384 kHz。传导 Ag-AgCl 头皮电极按照 International 10-20 System of Electrode Placement 标准放置。FFR 从激活的 Cz 电极记录下来，参考信号来自于两耳垂间电极。此外，用两个额外的电极监

测垂直眼动信号。记录中所有电极的偏置电压小于 20 mV。

### 3.4 实验程序

完成头皮电极放置后，受试者进入静音室 (single-walled Eckel C-14 booth, Cambridge, MA)，静坐在椅子上观看一部有字幕的无声电影，且不注意声源刺激。实验 1，每位受试者完成 1 组实验，实验过程约为 2 h。实验 1 的声源刺激在安静和噪声环境下分别重复播放 4000 次。数据分为 8 个模块，每个模块 1000 个试验。实验从安静环境开始，安静与噪声环境交替。每个模块内，正负刺激极性在试验之间交替进行，从而产生 2000 个重复的刺激 (2 种极性  $\times$  2 种环境)。实验 2 中，每位受试者在一天之内完成 4 个实验，每个实验过程约为 1 h。四个实验分别对应四个不同的声源刺激 (低频，中频，高频，宽频)，每个声源刺激在安静环境下重复 2000 次，正负极性交替进行。

## 4 实验结果与分析

### 4.1 时域细节结构 FFR 的 PLV 实验结果

图 1 显示 16 位受试者的平均  $\text{FFR}_C$  的 PLV，以及 PLV 显著有效的受试者人数。图 2 是在文献 [22] 的方法基础上计算的 16 位受试者的平均  $\text{FFR}_C$  的幅度谱 SNR，以及对应的显著有效的受试者人数。虽然幅度谱 SNR 已经能反映真实的 FFR 响应，但是，随着频率的增加， $\text{FFR}_C$  的 PLV 的有效人数多于幅度谱 SNR 的有效人数。因此，PLV 比幅度谱更有利分析  $\text{FFR}_C$ 。

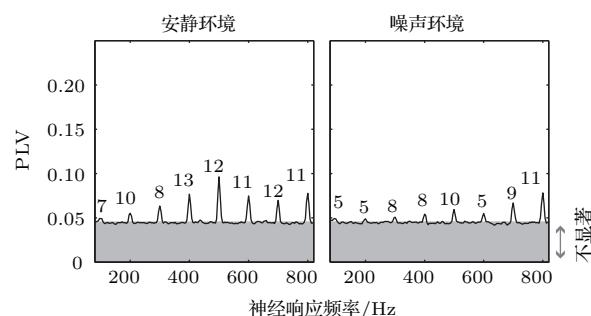


图 1  $\text{FFR}_C$  的平均 PLV 及显著有效人数  
Fig. 1. Mean PLVs of  $\text{FFR}_C$  and the number of subjects with significant efficient  $\text{FFR}_C$ .

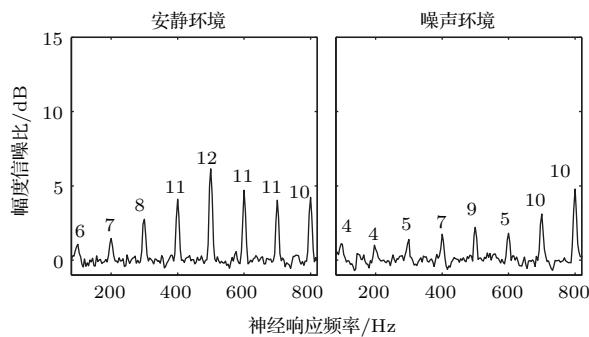
图 2 FFR<sub>C</sub> 的平均幅度谱 SNR 及显著有效人数

Fig. 2. Mean amplitude spectrum SNR of FFR<sub>C</sub> and the number of subjects with significant efficient FFR<sub>C</sub>.

## 4.2 样本熵实验结果

图 3(a) 是实验 1 刺激声为 100, …, 800 Hz 的复合音, 相似容限取 0.2, 维数是 2—11 时的 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的样本熵。离散点是每个受试者的样本

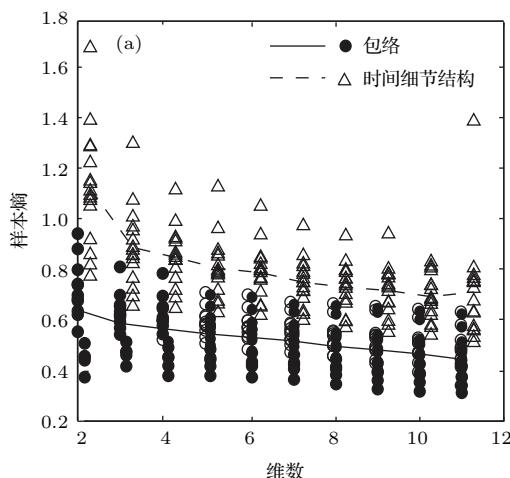
图 3 不同维数 (a)、不同频率 (b) 时 FFR<sub>E</sub> 与 FFR<sub>C</sub> 的样本熵比较 \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ 

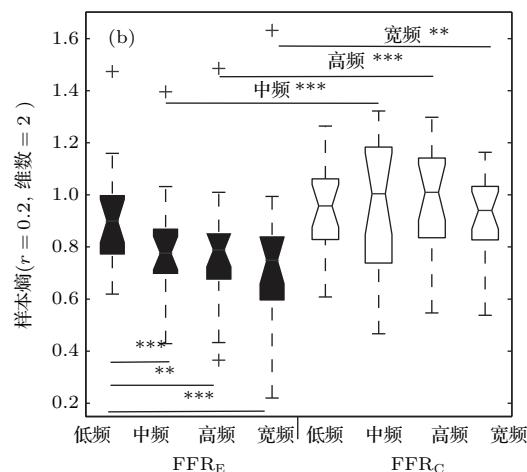
Fig. 3. Comparison between sample entropies of FFR<sub>E</sub> and FFR<sub>C</sub> under different dimensions (a) and frequencies (b); \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## 4.3 FFR<sub>ENV</sub> 与 FFR<sub>CAR</sub> 的正交独立关系

因为 PLV 的零假设是零均值高斯随机过程, 所以 PLV 的正交、独立与不相关三者等价。我们采用 Pearson's 线性相关系数表征 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的正交独立性。表 1 给出了安静环境中基频(100 Hz)、低频(200, 300, 400 Hz) 和高频(600, 700, 800 Hz) 三个频段内 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的 Pearson's 线性相关系数  $r$  以及显著水平  $p$ 。为保证计算准确性, 只有 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的 PLV 同时显著大于零假设的受试者才会被用于相关性计算, 表中给出了每组相关性的显著有效人数  $L$ 。用 Bonferroni 修正 15 组

熵, 曲线是样本熵均值。图 3(b) 是实验 2 刺激声为  $g_T(t)$  中频(600, …, 1100 Hz), 高频(1200, …, 1600 Hz) 和宽频(100, …, 2000 Hz), 相似容限取 0.2, 维数是 2 时 20 位受试者 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的样本熵, 凹线表示中值, 上、下边界分别表示第 25 个和第 75 个百分位数, 十字表示孤立点。

由图 3 可见: 1) 虽然增加维数会导致 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的样本熵逐渐减小, 但是各个维数对应的 FFR<sub>C</sub> 的样本熵均显著高于 FFR<sub>E</sub> 的样本熵, 实验 1 中两者样本熵在每个维数的差异水平  $p < 0.05$ ; 2) 实验 2 中维数为 2 时, 无论刺激声源频率是否可分辨, FFR<sub>C</sub> 的样本熵均显著高于 FFR<sub>E</sub> 的样本熵(中频,  $p < 0.001$ ; 高频,  $p < 0.001$ ; 宽频,  $p < 0.01$ )。由此可知, FFR<sub>E</sub> 和 FFR<sub>C</sub> 样本熵大的 FFR<sub>C</sub> 是对声音的快变成分 TFS 的响应, 而样本熵小的 FFR<sub>E</sub> 则是对声音的慢变成分包络的响应。



统计, 只有当显著水平  $p < 0.05/15 = 0.00333$  时 FFR<sub>E</sub> 和 FFR<sub>C</sub> 的线性相关才显著。显然, 任何频段 FFR<sub>E</sub> 和 FFR<sub>C</sub> 都不相关, 也即两个分量正交独立。噪声环境下二者也正交独立。所以, 本文提出的区分包络和时域细节结构响应的 PLV 方法在一定程度上解决了心理声学里很难分离包络和 TFS 这一问题。

## 4.4 声源基频处的 FFR<sub>E</sub> 和 FFR<sub>C</sub>

图 4 显示了在实验 2 的四种不同分辨率特性的复合音的刺激下,  $F_0$  处的 20 位受试者的平均 PLV (左, FFR<sub>E</sub>; 右, FFR<sub>C</sub>)。刺激声源包括可

分辨低频 100, …, 500 Hz (低频)、可分辨高频 600, …, 1000 Hz(中频)、不可分辨高频 1200, …, 1600 Hz (高频)、部分可分辨与部分不可分辨 100, …, 2000 Hz (宽频). 实验 2 的主要结果如下.

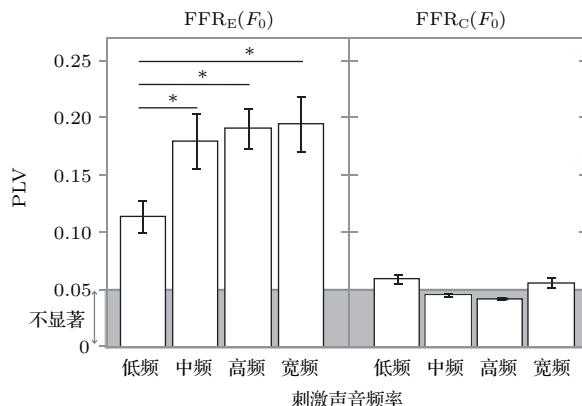


图 4 声源刺激频率影响基频的 FFR<sub>E</sub> 和 FFR<sub>C</sub> \*  $p < 0.05$   
Fig. 4. FFR<sub>E</sub> and FFR<sub>C</sub> at fundamental frequency are influenced by frequencies of sound stimuli, \*  $p < 0.05$ .

1) 以 PLV 为主要因素, 刺激声源为次要因素, 受试者作为误差做多因素 ANOVA 统计分析. 结果显示, 随着刺激声源所含的谐波阶次的提高, FFR<sub>E</sub> 在  $F_0$  的 PLV 有显著的增强, 并且刺激声音的频率和 PLV 有明显的交互作用 [ $F(3, 57) = 5.12$ ,

$p = 3.3 \times 10^{-3}$ ; 100—500 Hz,  $0.114 \pm 0.014$  SEM; 600—1000 Hz,  $0.179 \pm 0.024$  SEM; 1200—1600 Hz,  $0.191 \pm 0.018$  SEM; 100—2000 Hz,  $0.194 \pm 0.024$  SEM].

2) 不同的刺激声源之间两两做 ANOVA 统计分析, 结果表明: 在  $F_0$ , FFR<sub>E</sub> 对可分辨低频复合音低频的锁相强度明显低于其他复合音 [中频,  $F(1, 19) = 8.52$ ,  $p = 8.8 \times 10^{-3}$ ; 高频,  $F(1, 19) = 15.99$ ,  $p = 7.6 \times 10^{-4}$ ; 宽频,  $F(1, 19) = 9.95$ ,  $p = 5.2 \times 10^{-3}$ ]. 即使存在基频缺失时, FFR<sub>E</sub> 在基频仍然有响应. 而中频, 高频和宽频三种刺激声源两两之间的 PLV 并不存在显著差异 [中频 vs. 高频,  $F(1, 37) = 0.17$ ,  $p = 0.68$ ; 中频 vs. 宽频,  $F(1, 37) = 0.24$ ,  $p = 0.54$ ; 高频 vs. 宽频,  $F(1, 37) = 0.01$ ,  $p = 0.90$ ].

3) 与 FFR<sub>ENV</sub> 在  $F_0$  的强锁相响应不同的是, 只有当刺激声音含有  $F_0$  时, FFR<sub>C</sub> 在  $F_0$  的 PLV 才有锁相响应 [FFR<sub>C</sub> 在  $F_0$  的 PLV 均值是低频, 0.059; 宽频, 0.056]. 当刺激声源存在基频缺失时, FFR<sub>C</sub> 在  $F_0$  的 PLV 则接近零假设 [FFR<sub>C</sub> 在  $F_0$  的 PLV 均值是中频, 0.046; 高频, 0.042].

表 1 不同频段 FFR<sub>E</sub> 与 FFR<sub>C</sub> 的正交独立性

Table 1. Orthogonal and independent characteristics between FFR<sub>E</sub> and FFR<sub>C</sub> in different frequency bands.

		FFR <sub>E</sub>		FFR <sub>C</sub>	
FFR <sub>E</sub>	200—400	0.0074	-0.1755	-0.2602	-0.0626
	100	$p = 0.98$	$p = 0.65$	$p = 0.62$	$p = 0.85$
	$L = 16$	$L = 9$	$L = 6$	$L = 12$	$L = 12$
	600—800	0.6181	-0.2338	-0.5267	0.2047
	200—400	$p = 0.08$	$p = 0.66$	$p = 0.08$	$p = 0.21$
	$L = 9$	$L = 6$	$L = 12$	$L = 12$	
FFR <sub>C</sub>	200—400	-0.2726	-0.7584	-0.3566	
	600—800	$p = 0.66$	$p = 0.03$	$p = 0.35$	
	$L = 5$	$L = 8$	$L = 9$		
	100	0.5194	0.1801		
	200—400	$p = 0.48$	$p = 0.78$		
	$L = 4$	$L = 5$			
		0.3582			
		$L = 11$			

综上所述,  $F_0$  处的听觉响应主要是对声源包络的神经锁相编码, 而不是对 TFS 的编码. 在基频缺失的情况下,  $F_0$  处出现的畸变产物是不同的听觉通路对声源包络的混合响应, 而与 TFS 的无关.

宽频刺激声源(包含 100—2000 Hz 的全部谐波)会激励调谐到覆盖 20 个谐波的听觉外周通道, 其激励方式与对独立的低频、中频和高频的窄带输入的激励方式几乎相同. 若用一阶近似分析, FFR

对宽频刺激的响应应该是对各个窄带信号的响应的总和(加上文中没有给出其窄带信号 1100 Hz 和 1700—2000 Hz). 即使存在这样的叠加情况, 在任意给定的谐波点, 不同频率范围的相位的贡献也会不同, 相位的叠加可能使最终的响应增强, 也有可能使响应减弱. 事实上, 从实验结果我们可以推断出: 不同外周通道的包络响应的相位是不同的. 在基频  $F_0$ , 对宽频刺激的响应幅度小于对低频、中频和高频三个窄带刺激的响应的幅度总和.

#### 4.5 声源谐波处的 $\text{FFR}_E$ 和 $\text{FFR}_C$

上一节的分析表明无论刺激声音中基频  $F_0$  是否缺失,  $F_0$  处的  $\text{FFR}_E$  的神经元锁相响应是所有频率点中最强的. 不同的 FFR 分量的相位同步能力不仅在基频不同, 而且在其他的谐波点之间也不相同. 实验 1 的声源谐波可分为低频可分辨频段(200, 300 和 400 Hz)、中频可分辨频段(500 Hz)以及高频可分辨频段(600, 700 和 800 Hz). 安静环境中, 有以下结论:

- 1)  $\text{FFR}_E$  的 PLV 集中于低频可分辨频段, 且其 PLV 强于中频和高频;
- 2) 随着可分辨谐波阶次的增加,  $\text{FFR}_E$  的 PLV 随之降低 [ $F(2, 30) = 34.06, p = 1.9 \times 10^{-8}$ ; 200—400 Hz,  $0.110 \pm 0.010$  SEM; 500 Hz,  $0.080 \pm 0.009$  SEM; 600—800 Hz,  $0.055 \pm 0.004$  SEM];
- 3) 与  $\text{FFR}_E$  相反,  $\text{FFR}_C$  的 PLV 则集中于中、高频可分辨频段;
- 4)  $\text{FFR}_C$  的 PLV 随着频率的增加而显著提高 [ $F(2, 30) = 4.66, p = 1.7 \times 10^{-3}$ ; 200—400 Hz,  $0.059 \pm 0.004$  SEM; 500 Hz,  $0.076 \pm 0.008$  SEM; 600—800 Hz,  $0.063 \pm 0.006$  SEM];

本文进一步分析了不同频段、不同刺激声音的 FFR 的四种 PLV, 即 100—500 Hz 的  $\text{FFR}_E$ , 100—500 Hz, 600—1000 Hz 和 1200—1600 Hz 的  $\text{FFR}_C$ , 如图 5 所示. 声源的频率听觉对包络和时域细节结构的编码有如下影响.

- 1) 除包含各个频率的 100—2000 Hz 宽带声音外, 在每个频段, 都是与频段相同的刺激声音的  $\text{FFR}_C$  最强, 也即  $\text{FFR}_C$  对刺激频率本身的锁相能力强于对其他频率的锁相(各频段最强的  $\text{FFR}_C$ : 100—500 Hz,  $0.124 \pm 0.013$  SEM; 600—1000 Hz,  $0.170 \pm 0.019$  SEM; 1200—1600 Hz,  $0.273 \pm 0.031$  SEM; 100—2000 Hz,  $0.240 \pm 0.033$  SEM).

2) 当刺激声音是不可分辨谐波 1200—1600 Hz 时, 虽然各个试验之间对包络的响应之间不同相, 但是, 此时  $\text{FFR}_C$  的 PLV 却显著高于两个可分辨复合音 100—500 Hz 和 600—1000 Hz 的  $\text{FFR}_C$  的 PLV[高频 vs. 低频,  $F(1, 19) = 30.20, p = 2.7 \times 10^{-5}$ ; 高频 vs. 中频,  $F(1, 19) = 37.67, p = 6.7 \times 10^{-6}$ ].

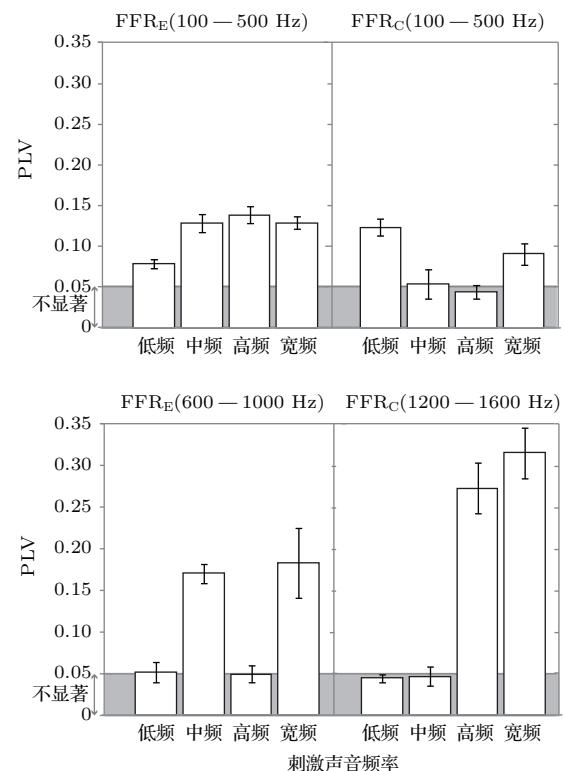


图 5 刺激声源的频率特性对各个频段 FFR 的影响

Fig. 5. FFRs at different frequency bands are influenced by frequencies of sound stimuli.

3) 在 1200—1600 Hz, 尽管 100—1000 Hz 包含多个不可分辨谐波, 但是其  $\text{FFR}_C$  的 PLV 并不明显地高于 1200—1600 Hz 刺激对应的响应 [ $F(1, 19) = 3.27, p = 8.0 \times 10^{-2}$ ].

4) 由于 100—2000 Hz 刺激包含了全部的可分辨和不可分辨频率, 因此对应的 PLV 高于只包含可分辨频率的两个刺激声源(100—500 Hz 和 600—1000 Hz)的  $\text{FFR}_C$  的 PLV[100—2000 Hz,  $0.240 \pm 0.033$  SEM; 100—500 Hz,  $F(1, 19) = 18.52, p = 3.8 \times 10^{-4}$ ; 600—1000 Hz,  $F(1, 19) = 17.61, p = 4.9 \times 10^{-4}$ ].

5) 不仅不可分辨复合音产生的  $\text{FFR}_C$  的 PLV 高于可分辨谐波, 而且可分辨复合音之间在对应的频率处也有差异. 中频刺激声音(600—1000 Hz, 基频缺失)在中频频段的锁相能力强于低频刺激声音

(100—500 Hz, 包含基频)在低频频段的锁相能力 [ $F(1, 19) = 8.13, p = 1.0 \times 10^{-2}$ ].

## 5 结 论

为了分析人的听觉神经系统对声源的包络和时间细节结构的神经元锁相机理, 本文在文献[22]的基础上提出基于样本熵的时域细节结构频率跟随响应的锁相值分析方法. 听觉认知脑电实验和PLV数据分析验证了新方法的有效性. FFR<sub>T</sub>的PLV样本熵显著大于FFR<sub>E</sub>的PLV样本熵, 二者正交独立. 因此, 新方法能够分别反映听觉系统对声音不同分量的锁相能力. 且该方法计算量小, 有效人数多, 不受 $1/f$ 多及工频噪声影响.

本文发现听觉神经主要锁相机理: 1) 基频处的FFR<sub>E</sub>大于谐波处响应, 但是未见明显的FFR<sub>T</sub>, 表明基频处的响应主要来源于神经元对包络的锁相; 2) 在基频处FFR对不可分辨谐波成分包络的锁相能力强于其对可分辨谐波成分包络的锁相能力; 3) 在基频缺失的情况下, 畸变产物来源于不同的听觉通路的FFR<sub>E</sub>的混合作用, 该结论与已有的音高感知等认知和心理实验结论一致[22–24], 进一步说明了方法的有效性, 由此可推测听觉神经通路对基频的编码可能取决于基频、包络整流、谐波的整流以及耳蜗畸变产物; 4) 在谐波处, FFR<sub>E</sub>集中于慢变的低频, 而FFR<sub>T</sub>则集中于快变的中、高频; 5) FFR<sub>E</sub>与FFR<sub>T</sub>的神经元锁相能力与声源频率的可分辨性相关. 在以上结论的基础上, 我们下一步工作是将脑皮层下听觉系统对声音包络和TFS的PLV映射到耳蜗定位和神经元定时神经编码机理, 从而进一步研究在复杂环境下人的听觉系统的神经元放电机理.

感谢Center for Computational Neuroscience and Neural Technology, Boston University的Barbara G Shinn-Cunningham教授和Hari Bharadwaj博士, 以及清华大学生物医学工程系的胡广书教授对本文的实验设计及方法的指导.

## 参考文献

- [1] Moore B C J, Patterson R D, Winter I M, Carlyon R P, Gockel H E 2013 *Basic Aspects of Hearing: Physiology and Perception* (New York: Springer) pp12–25
- [2] Shen B K, Wang J F, Zeng T 2006 *Chin. Phys. Lett.* **23** 3380
- [3] Thatcher R W 2012 *Dev. Neuropsychol.* **37** 476
- [4] Yang L X, Chen K A, Zhang R B, Liang Y 2014 *Acta Phys. Sin.* **63** 134304 (in Chinese) [杨立学, 陈克安, 张瑞冰, 梁雍 2014 物理学报 **63** 134304]
- [5] Du Y, Buchsbaum B R, Grady C L, Alain C 2014 *Proc. Natl. Acad. Sci. USA* **111** 7126
- [6] Lehmann A, Schönwiesner M 2014 *PloS one* **9** e85442
- [7] Ruggles D, Bharadwaj H, Shinn-Cunningham B G 2012 *Curr. Biol.* **22** 1417
- [8] Wu X B, Mo J, Yang M H, Zheng Q H, Gu H G, Ren W 2008 *Chin. Phys. Lett.* **25** 2799
- [9] Huang X L, Huo C Y, Si J F, Liu H X 2014 *Acta Phys. Sin.* **63** 100503 (in Chinese) [黄晓林, 霍铖宇, 司峻峰, 刘红星 2014 物理学报 **63** 100503]
- [10] Peeters G, Giordano B L, Susini P, Misdariis N, MaAdams S 2011 *J. Acoust. Soc. Am.* **130** 2902
- [11] Ali M S, Saravanakumar R 2015 *Chin. Phys. B* **24** 050201
- [12] Wang J S, Wang M L, Li X L, Niebur E 2015 *Chin. Phys. B* **24** 038701
- [13] Qin F, Zhang Q X, Deng X H 2012 *Chin. Phys. B* **21** 040701
- [14] Ding X L, Li Y Y 2014 *Acta Phys. Sin.* **63** 248701 (in Chinese) [丁学利, 李玉叶 2014 物理学报 **63** 248701]
- [15] Plack C J, Oxenham A J, Popper A N, Fay R 2005 *Pitch* (New York: Springer) pp169–233
- [16] Moore B C 2008 *J. Assoc. Res. Oto.* **9** 399
- [17] Ruggles D, Bharadwaj H, Shinn-Cunningham B G 2011 *Proc. Natl. Acad. Sci. USA* **108** 15516
- [18] Zeng F G, Nie K, Stickney G S, Kong Y Y, Vongphoe M, Bhargave A, Wei C, Cao K 2008 *Proc. Natl. Acad. Sci. USA* **105** 2293
- [19] Smith Z M, Delgutte B, Oxenham A J 2002 *Nature* **416** 87
- [20] Zhu L 2013 *Ph. D. Dissertation* (Beijing: Tsinghua University) (in Chinese) [朱莉 2013 博士学位论文 (北京: 清华大学)]
- [21] Hopkins K, Moore B C 2009 *J. Acoust. Soc. Am.* **125** 442
- [22] Zhu L, Bharadwaj H, Xia J, Shinn-Cunningham B G 2013 *J. Acoust. Soc. Am.* **134** 384
- [23] Oxenham A J, Michey C, Keebler M V 2009 *J. Acoust. Soc. Am.* **125** 2189
- [24] Brown C A, Bacon S P 2010 *Hear. Res.* **266** 52

# Experimental analysis of auditory mechanism of neural phase-locking based on sample entropy\*

Zhu Li<sup>1)</sup> Deng Juan<sup>2)</sup> Wu Jian-Hua<sup>1)</sup> Zhou Nan-Run<sup>1)</sup>†

1) (School of Information Engineering, Nanchang University, Nanchang 330031, China)

2) (Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China)

(Received 12 March 2015; revised manuscript received 23 April 2015)

## Abstract

Phase-locking is a physical phenomenon that refers to a system response which is synchronized with a specific phase of the periodic stimulus. The auditory neural phase-locking plays an important role in revealing the basic neural mechanism of auditory cognition and improving auditory perception. In the existing auditory researches, psychophysical and amplitude spectral methods are mainly adopted. However, those two methods cannot differentiate the envelope-related auditory response from the temporal-fine-structure-related auditory response, and cannot reveal the neural phase-locking mechanism directly either. In this study, a phase locking value (PLV), based on sample entropy, bootstrapping and discrete Fourier transform, is proposed for analyzing the temporal-fine-structure-related frequency following response (FFRT). The proposed PLV is applied to computing neural and physical data. Two electroencephalography experiments are carried out. Results show that the sample entropy of FFRT's PLV is significantly greater than that of FFR<sub>E</sub>'s PLV, and the two PLVs are orthogonal and independent. Thus, the PLVs of FFR<sub>E</sub> and FFRT reveal the auditory phase-locking mechanisms effectively. In addition, the response to fundamental frequency is mainly attributed to the envelope-related phase locking. And human auditory capability of phase locking to the envelope of the unresolved frequency is superior to the capability of phase-locking to the envelope of the resolved frequency. Moreover, in the case of missing fundamental frequency, the distortion product is the mixture of FFR<sub>E</sub> in various auditory neural paths. Also, FFR<sub>E</sub> concentrates at the low harmonic frequencies, while FFRT concentrates at the mid and high order harmonic frequencies. Therefore, the auditory neural phase-locking is related to the frequency resolution of sound. In conclusion, the proposed method overcomes some disadvantages of existing FFR analyses, making it beneficial to exploring auditory neural mechanisms.

**Keywords:** neural mechanism of audition, phase locking value, sample entropy, frequency following response

**PACS:** 43.64.-q, 43.60.Np, 89.70.Cf

**DOI:** [10.7498/aps.64.184302](https://doi.org/10.7498/aps.64.184302)

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 61463035, 61262084), the Natural Science Foundation of Jiangxi Province, China (Grant Nos. 20142BAB217022, 20122BAB211020), and the Natural Science Foundation of Jiangxi Education Commission, China (Grant No. GJJ14193).

† Corresponding author. E-mail: [znr21@163.com](mailto:znr21@163.com)