对……的阐释与复杂序列结构的稳态神经反应分析：方法学说明

丁鼎

生物医学工程与仪器科学学院，

浙江大学，中国杭州

摘要

频率标记是一种强大的方法，可用于研究感觉特征的神经加工。近年来，这一方法也被用于研究复杂序列（如言语和音乐）中上位结构（即组块）的神经相关物。序列结构的嵌套、控制感觉特征周期性的必要性，以及序列结构的低频特性，都给数据分析和解释带来了新的挑战。本文讨论如何解释序列结构的频率，以及分析信号周期性时需要考虑的因素。最后，本文推荐一种稳妥的频率标记反应分析流程。

1. 引言

频率标记是一种强大的技术，可用于提取跟踪刺激特征的神经反应。一般而言，在频率标记范式中，目标刺激特征会以频率 f 周期性调制。因此，动态跟踪该刺激特征的神经反应也会以频率 f 波动。通常使用离散傅里叶变换或小波变换提取 f Hz 的频率标记反应。频率标记是脑电图（EEG）和脑磁图（MEG）研究中的一种强大范式，因为无论反应的潜伏期或波形如何，它都能提取任何跟随刺激中 f Hz 变化的神经反应。该范式已广泛应用于视觉 （Norcia 等，2015；Regan，1977）和听觉（Galambos 等，1981；Picton 等，2003）加工研究：对视觉特征（如亮度）的周期性变化所产生的频率标记反应称为稳态视觉诱发电位（SSVEP）；对听觉特征（如强度）的周期性变化所产生的频率标记反应则称为听觉稳态反应（aSSR）。这些反应广泛用于研究感觉编码的基本特性（Herrmann，2001；Ross 等，2000；Wang 等，2012；Wong 等，2007）和认知控制（Andersen 等，2008；Elhilali 等，2009；Gao 等，2021）。

近年来，频率标记范式已被用于研究复杂序列中上位结构的神经加工，例如言语和音乐：这些研究的假设是，心理构建的上位序列结构，即一个句子，会通过神经响应得到表征，而该响应的持续时间与刺激中的结构时长相匹配（Buiatti 等，2009；Ding 等，2016；Nozaradan 等，2011）。一方面，频率标记提供了一种

强大的范式，可用于研究组块的神经加工，并将其与短暂的刺激事件进行对比；这一范式也催生了大量研究（Batterink 与 Paller，2019；Benitez-Burraco 与 Murphy，2019；Choi 等，2020；Glushko 等，2022；Henin 等，2021；Kaufeld 等，2020；Kazanina 与 Tavano，2022；Keitel 等，2018；Lo 等，2022；Lu 等，2021；Makov 等，2017；Meng 等，2021；Meyer，2018）。另一方面，序列加工问题的复杂性也给频率标记反应的分析和解释带来了更多挑战。首先，在传统的频率标记研究中，每个感兴趣的刺激特征都会被标记为一个独特的频率；而复杂序列中的结构往往是嵌套的，因此不同层级的结构无法被标记为彼此无关、任意设定的频率。例如，在句子“the cute boy s miled”中，前三个词根据句法构成一个名词短语。然而，三词名词短语与四词句子彼此嵌套，因此无法被标记为彼此无关的频率。结构之间的嵌套导致结构时长与结构重复周期相分离，这一点将在第2.1节中讨论。

其次，传统的频率标记研究会明确地使刺激特征发生周期性变化，而序列结构研究有时则希望避免基本刺激特征出现此类周期性变化，以分离由内部心理过程产生的神经反应。什么是由内部心理过程产生的神经反应？例如，聆听等时节拍序列时，人们可能会在心中想象出一种韵律结构，而想象中的节拍速率处的神经反应可以反映由内部驱动的过程（Nozaradan 等，2011）。类似地，当一串词语依据句法规则被组合成句子时，句子速率处的神经反应可以反映高层次句子加工（Ding 等，2016）。然而，在这些情况下，如果某种基本感觉特征与想象中的节拍或

句法构建的句子具有相同的周期性，就无法确定神经反应反映的是感觉特征还是序列结构。因此，通常有必要检查刺激特征的周期性。但需谨慎，因为某些类型的周期性无法通过傅里叶变换捕捉到，第2.2节将对此进行讨论。

第三，对频率标记序列结构反应的分析有时容易受到伪迹影响，而这类伪迹很少影响传统频率标记反应的分析。序列结构通常对应非常低的频率，例如 < 3 Hz，而如此低的频率可能受到分析时段重叠的污染（Benjamin 等，2021）。第3节将说明这类伪迹为何会产生，并讨论恰当分析频率标记反应的潜在指南，包括如何选择分析时长，以及是否应使用平滑窗。本文讨论的是常见的技术问题，而非某项特定实验的分析。不过，为便于理解，本文在图1A中提供了一个假设实验，但所得结论并不限于此例。另一方面，本文的目标读者是实验研究人员，而非工程师。因此，本文尝试通过图示阐释相关概念，并省略数学推导。离散傅里叶变换的数学基础可参见 Oppenheim 等（2001）等经典教材。

2. 傅里叶变换未能反映什么

2.1 频率可能无法反映时间常数或信号持续时间

对于频域分析而言，频率是一个核心概念，它与信号周期相对应。然而，信号周期不一定与信号的其他时间常数一致。例如，指数信号 $e^{-t/\tau}$ 具有

时间常数τ，但该信号是非周期性的，τ 并非信号的周期。即使对于周期信号，其周期也可能与周期内波形的时间常数或持续时间不同，图1B展示了一些例子。在这些例子中，信号的周期为 1 s，左图展示了一个周期内的波形。信号的时间参数，包括指数函数的时间常数、锯齿波信号的持续时间，以及单周期正弦波的频率，会影响傅里叶频谱的形状，但通常不会产生与这些参数对应的频谱峰值。相反，由于信号以 1 Hz 的频率重复，频谱会在 1 Hz 及其谐波相关频率处出现峰值，即 2 Hz、3 Hz 等。然而，当信号周期发生变化时，即使周期内的波形保持不变，谱峰也会相应移动（图1C）。关于信号重复率和周期内波形如何影响频谱的更多示例，请参见 Zhou 等人（Zhou et al., 2016）。

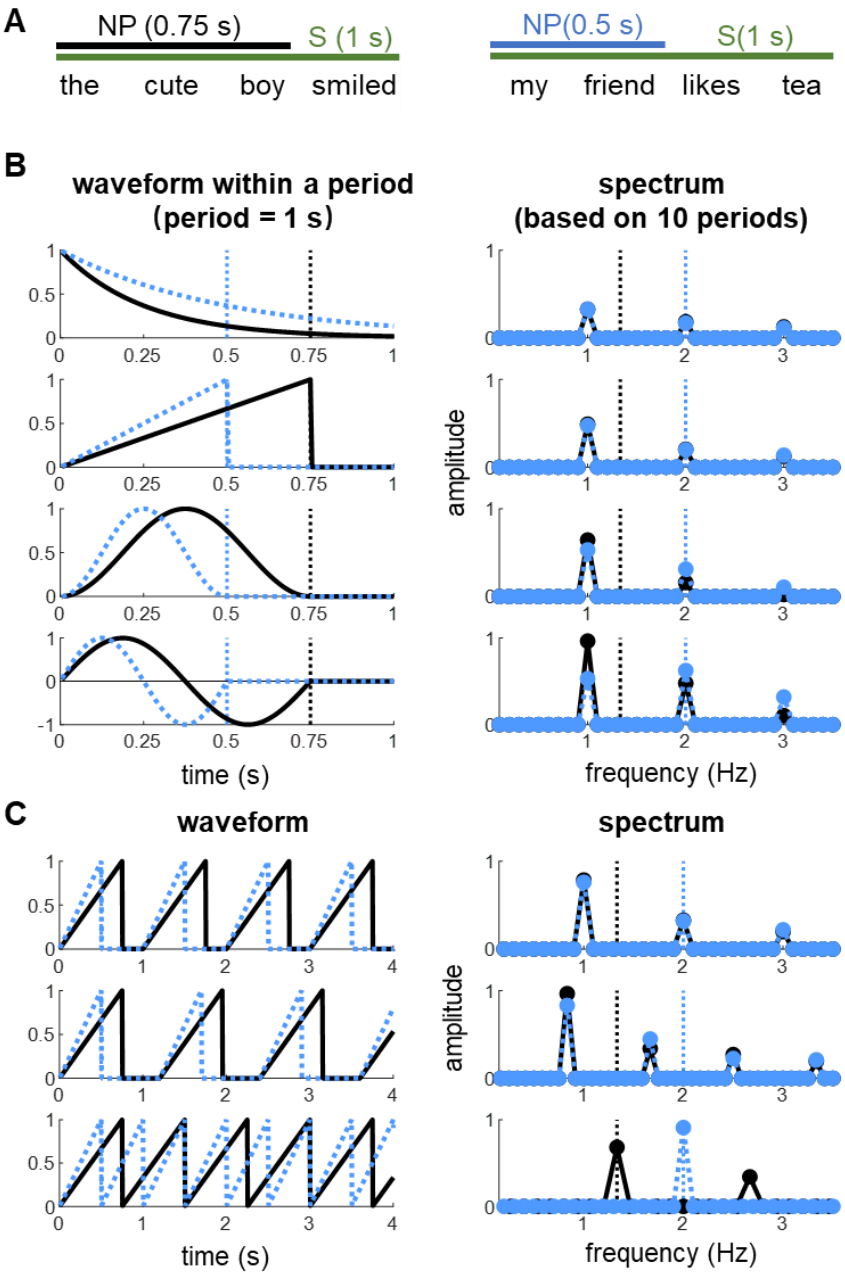


图 1. 频谱中的峰值反映了信号的周期性。A）一种假设的实验条件，其中一个名词短语（NP）嵌入在一个句子（S）中。NP 的持续时间为 0.75秒或 0.5秒，并假设神经反应受 NP 持续时间的调制。B）每隔 1秒重复一次的信号及其对应的频谱。左图显示一个周期内的波形，黑色和蓝色曲线的时间常数不同，分别为 0.75秒和 0.5秒。右图显示频谱，即对应信号 10 个周期的离散傅里叶变换的幅值。时间常数及其对应的频率以垂直虚线标示。无论信号在一个周期内的时间常数如何，频谱都在 1 Hz（即信号周期的倒数）及谐波相关频率处出现峰值。C）由相同锯齿波波形构成、但重复频率不同的信号。谱峰始终反映重复频率。

1.2 频率可能无法反映变化率

假设某个信号每隔 T 秒变化一次。直观来看，其傅里叶频谱应在 $1/T$ Hz 处达到峰值。然而，这种直觉并不总是正确，图 2 给出了一个例子：频谱在 $1/T$ Hz 及其谐波相关频率处呈现波谷。当该信号用于调制一个 4 Hz 正弦波的增益时，调制后的正弦波在 $1/T$ Hz 处同样没有功率。这些例子旨在说明，傅里叶变换可能无法检测到某些节律。为什么该信号在 $1/T$ Hz 处没有功率？在傅里叶变换中，f 处的功率由信号与频率为 f 的正弦波（包括正弦和余弦）之间的点积决定。图 2 中的信号在每个 T 秒内都没有波动，因此该信号与 $1/T$ Hz 的正弦波没有相关性。图3展示了信号之间的点积。

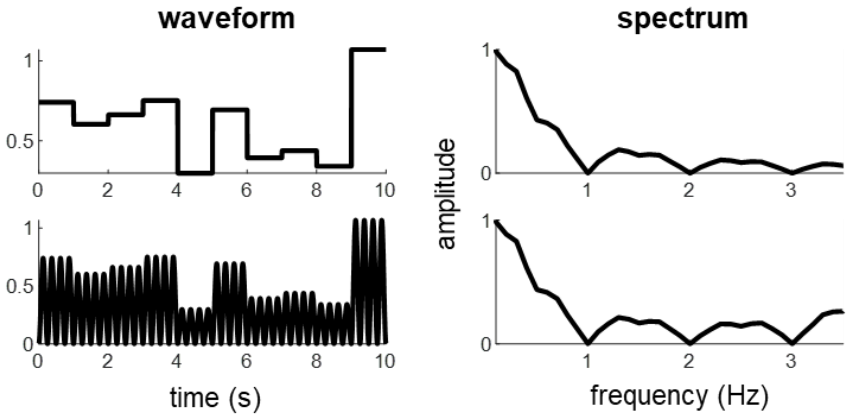


图 2. 信号的变化率可能对应频谱中的波谷。上图显示一个每 1 秒变化一次的信号，下图则是一个 4 Hz 的正弦波，其幅度由上图中的信号调制。在频谱中，可以观察到 1 Hz 及谐波相关频率处的波谷。

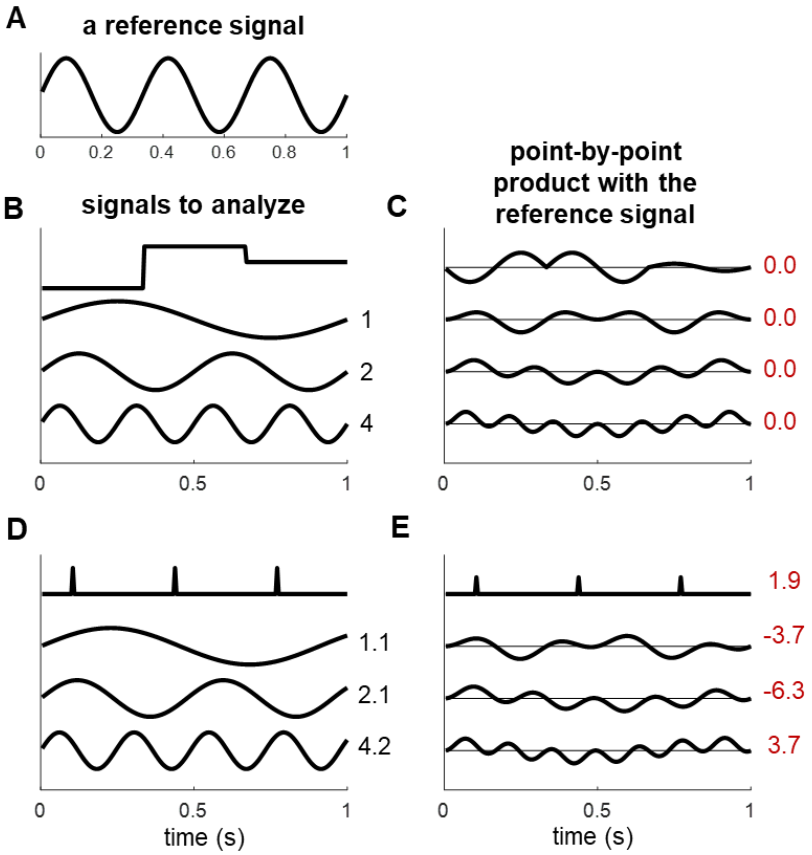


图3. 信号之间点积的示意图，这是离散傅里叶变换的基础。A）一个 3 Hz 的正弦波，用于计算 3 Hz 处的离散傅里叶变换系数。BC）待分析的信号及其与参考信号逐点相乘的结果。顶部信号与图 2 中的信号相似，另外 3 个信号则是正弦波，其频率由 B 面板中的数字标示。乘积信号的总和（即两个信号之间的点积）由 C 面板中的红色数字标示。DE）与参考信号的点积不为零的信号示例。

2. 神经反应分析方法的影响

2.1 重叠时间段可能引入伪迹

通过将信号的延迟版本与其自身相加，可以基于任意信号创建一种节律。图 4A 展示了一个例子：待分析的信号仅在 4.8 秒处包含一个脉冲，其余时间均为 0。然而，当信号被划分为 5 秒的时间段，且相邻时间段重叠 4 秒时，平均时间段显然会变为周期性，其周期与相邻时间段之间的间隔相同，例如 1 秒。另一个

示例如图所示。4B 中，白噪声以同样的方式被切分为 5秒 时间段。对 100 个时间段求平均后的频谱清晰地显示出 1 Hz 的峰值。事实上，在听觉研究中，这种方法已被用于基于白噪声等信号产生音高感知（Yost, 1996）。

如果不恰当的数据分段会引入伪迹，为什么不直接对未分段的数据应用傅里叶变换？直接对未分段的数据进行傅里叶变换，确实可以得到响应的高频率分辨率频谱。然而，在实际的脑电图/脑磁图记录中，长时间记录期间几乎无法避免由头部运动或硬件故障等造成的强伪迹；在脑电图/脑磁图分析中，排除存在大量伪迹的记录是常见做法。不过，仅仅因为少数分散出现的伪迹就丢弃一段长时间记录并不理想。因此，将长时间记录切分成较短的时间段，并仅剔除存在明显伪迹的时间段，是一种常见策略

2.2 分析窗决定谱峰的宽度

假设频率标记神经反应的周期为T秒，并将D秒的记录数据通过离散傅里叶变换转换到频域。离散傅里叶变换频谱由对应离散频率的系数组成，即 1/D Hz、2/D Hz、3/D Hz 等。如果D是T的整数倍，频谱中就能分辨出频率标记响应。换言之，如果$D = kT$，其中k为整数，那么第k^{th} 个离散傅里叶变换系数对应的频率为 1/T Hz，即目标频率。在这种情况下，响应频谱仅在 1/T Hz 及其谐波相关频率处具有功率。图5A（上图）展示了一个例子，其中T 为 0.5 秒，D 为 5 秒，神经反应恰好是一个正弦波。响应频谱在 4 Hz 处有一个尖峰，且功率

相邻频率箱中的功率为 0。不在 4 Hz 处的离散傅里叶变换系数均为零，因为频率为离散傅里叶变换可分辨频率的任意两个、时长均为 D 秒的正弦波之间的点积为零（图3BC）。然而，当 D 不是 T 的整数倍时，离散傅里叶变换频谱中没有对应于 $1/T$ Hz 的频率箱，信号的功率会扩散到 $1/T$ Hz 附近的多个频率箱中，这种现象称为频谱泄漏。图5B（上图）展示了一个例子，其中 T 仍为 0.5 秒，而 D 为 5.1秒。

.

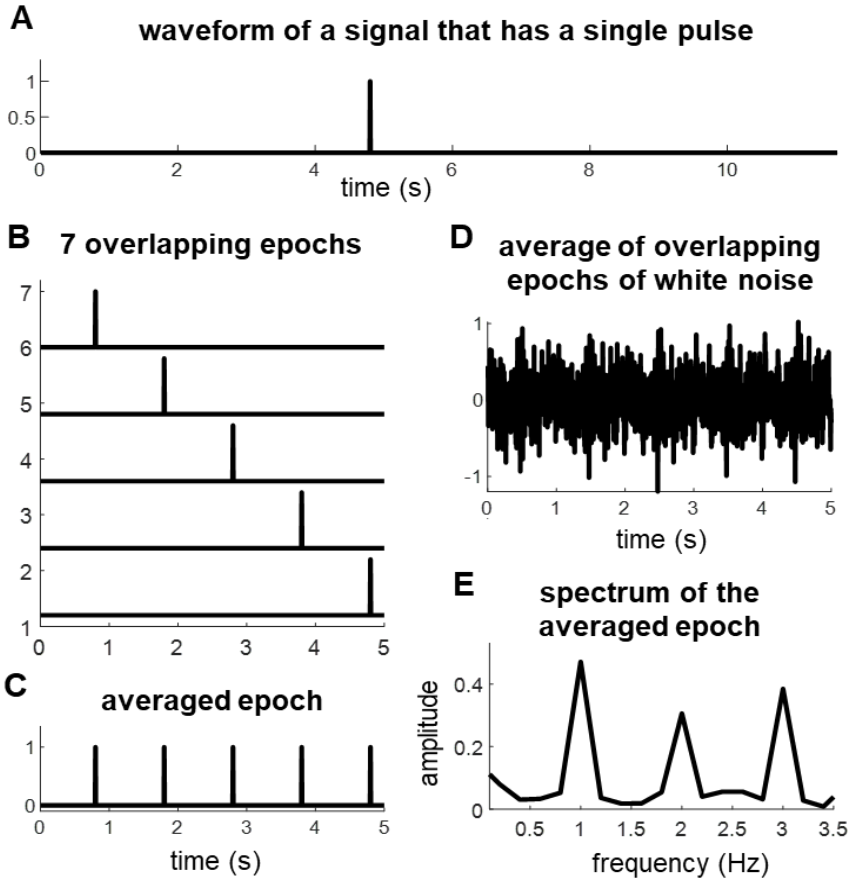


图4. 重叠时间段可能导致频谱中出现伪峰。A）由单个脉冲组成的非周期信号。B）将A中的信号分割为5秒时间段，彼此重叠4秒。C）对B中的时间段取平均。D）对白噪声应用相同的时间段划分过程，并显示所得波形。E）D中信号的频谱。为获得稳健的结果，生成20组相互独立的白噪声，并以相同方式处理，然后对频谱取平均。

缓解频谱泄漏的一种常见策略是在进行傅里叶变换前，将平滑窗乘到信号上。加窗信号的频谱如图5（下图）所示。使用平滑窗后，信号持续时间不再显著影响频谱的形状，但目标频率（即 2 Hz）附近的频率箱中始终存在非零功率。图5中各种方法的主要区别在于，正弦波的全部功率是集中在单个频率箱中，还是分散到多个频率箱中。本文不再进一步说明，但这些结论也适用于分析方法的其他变体，例如在信号中补零，或使用小波变换而非傅里叶变换。

我们是否需要关注信号功率是否集中在单个频率箱中？在某些情况下，答案是肯定的。例如，检验频率标记反应统计显著性的一种简便方法，是将目标频率处的功率与相邻频率箱中的功率进行比较（Benjamin 等，2021；Ding 等，2016；Nozaradan 等，2011）。当相邻频率箱中的功率高于基线时，这种方法的统计功效显然会受到削弱。即使采用其他方法检验频率标记响应的统计显著性，例如将其与不包含频率标记响应的控制条件进行比较（Andersen 等，2008），若能将频率标记响应的全部功率集中在单一频率上，检验的统计功效也会有所提升。

更一般地说，当信号的周期性未知、需要通过傅里叶分析来确定时，平滑窗往往会有所帮助。然而，在频率标记范式中，目标频率是已知的，因此

平滑窗并非必需。换言之，在频率标记方法中，数据分析的目的不是估计响应的周期性，而是检测是否存在具有已知频率的响应。根据信号检测理论（Poor，1998年），要对正弦波进行最优检测，通常需要计算记录到的信号与目标信号之间的点积；在频率标记范式中，目标信号可视为正弦波，而点积可通过离散傅里叶变换方便地计算。

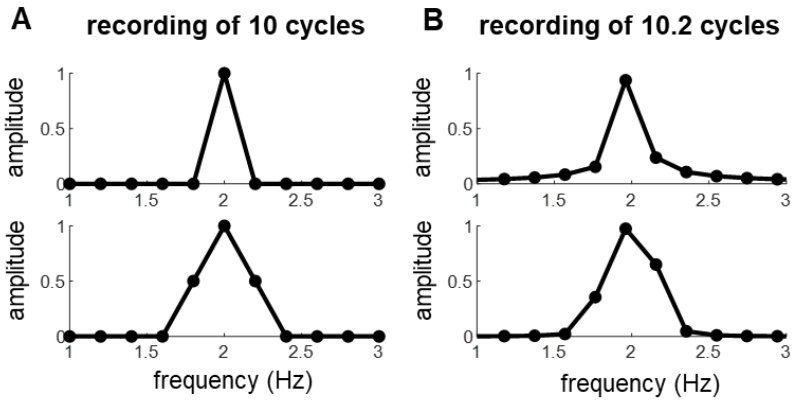


图5. 频谱泄漏与加窗。待分析的信号是一个2赫兹正弦波，A图中的信号持续时间为5秒，B图中为5.1秒。上图为信号的离散傅里叶变换，下图为经汉宁窗平滑后的信号的离散傅里叶变换。

3. 总结

首先，在频率标记范式中，目标频率是刺激特征或序列结构重复出现的频率，通常与该特征或结构持续的时间或其在每个周期内变化的快慢无关。其次，傅里叶变换并不能提供一种放之四海而皆准的方法来提取信号中的所有周期性。在刺激方面，需要谨慎处理，例如确保刺激不包含任何可能出现在目标频率上的周期性。在响应方面，可能需要更先进的特征提取方法来识别频率标记响应。例如，对于

图 2 中的信号，取信号一阶导数的绝对值可以反映信号中的 1 赫兹周期性。

最后，我建议采用以下相对稳妥的流程来分析频率标记反应。

(1) 待分析的响应应恰好包含频率标记响应的整数个周期。更具体地说，如果响应采样率为F，且响应以f进行频率标记，则每个响应周期的采样点数为F/f，该数值不必为整数。不过，如果分析窗中包含k个周期，那么分析窗的总长度，即kF/f，应为整数，且k也应为整数。

(2) 当刺激持续时间很长（例如数分钟）时，可将刺激呈现过程中记录的响应直接转换到频域。另一种做法是将其分割成较短的时间段（例如，剔除伪迹较大的时间段），再进行平均。不过，这些时间段不应重叠。

(3) 当目标频率已知，且分析窗包含目标响应的整数个周期时，进行傅里叶分析无需使用平滑窗。

致谢

感谢 Wenhui Sun 帮忙整理参考文献。本研究得到国家自然科学基金委员会（32222035）和浙江省重点研发计划（2022C03011）的资助。

参考文献

Andersen, S. K., Hillyard, S. A., & Muller, M. M. (2008年). 注意可促进人类视觉皮层并行处理多种刺激特征。《Curr Biol》，18(13), 1006-1009. https://doi.org/10.1016/j.cub.2008.06.030

Batterink, L. J., & Paller, K. A. (2019). 即使不在注意焦点内，也能进行语音规律的统计学习。《Cortex》，115, 56-71. https://doi.org/10.1016/j.cortex.2019.01.013

Benitez-Burraco, A., & Murphy, E. (2019). 为什么脑振荡正在增进我们对语言的理解。《Front Behav Neurosci》，13, 190. https://doi.org/10.3389/fnbeh.2019.00190

Benjamin, L., Dehaene-Lambertz, G., & Fló, A. (2021). 关于稳态反应分析的评述：重叠时间段引入的伪影。《Cortex》，142. https://doi.org/10.1016/j.cortex.2021.05.023

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). 通过频率标记神经电反应探究连续语音计算的神经相关物。神经影像学, 44(2), 509-519. https://doi.org/https://doi.org/10.1016/j.neuroimage.2008.09.015

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). 言语前婴儿发现词语统计模式的速度与成年人相近：来自神经同步的证据。心理科学, 31(9), 1161-1173. https://doi.org/10.1177/0956797620933237

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). 连续语音中层级语言结构的皮层追踪。Nat Neurosci, 19(1), 158-164. https://doi.org/10.1038/nn.4186

Elhilali, M.、Xiang, J.、Shamma, S. A. 和 Simon, J. Z.（2009）。注意与自下而上的显著性之间的交互

介导了听觉场景中前景和背景的表征。《公共科学图书馆·生物学》，7(6)，e1000129。https://doi.org/10.1371/journal.pbio.1000129

Galambos, R., Makeig, S., & Talmachoff, P. J. (1981). 从人类头皮记录的40赫兹听觉电位。美国国家科学院院刊, 78(4), 2643-2647. https://doi.org/10.1073/pnas.78.4.2643

Gao, X., Wang, Y., Chen, X., & Gao, S. (2021). 广义脑–计算机接口中的接口、交互与智能。《认知科学趋势》, 25(8), 671-684. https://doi.org/10.1016/j.tics.2021.04.003

Glushko, A., Poeppel, D., & Steinhauer, K. (2022). 显性韵律和隐性韵律会影响此前被归因于语法加工的神经生理反应。《科学报告》，12(1), 14759. https://doi.org/10.1038/s41598-022-18162-3

Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle, W., Devinsky, O., & Melloni, L. (2021). 在人类大脑皮层和海马体中学习层级序列表征。《科学进展》，7(8), eabc4530. https://doi.org/10.1126/sciadv.abc4530

Herrmann, C. S. (2001). 人类脑电图对 1-100 Hz 闪烁的反应：视觉皮层中的共振现象及其与认知现象的潜在相关性。Exp Brain Res, 137(3-4), 346-353. https://doi.org/10.1007/s002210100682

Kaufeld, G., Bosker, H. R., Ten Oever, S., Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). 语言结构和意义将神经振荡组织成内容特异性层级。《神经科学杂志》，40(49), 9467-9475. https://doi.org/10.1523/JNEUROSCI.0302-20.2020

Kazanina, N., & Tavano, A. (2022). 神经振荡对句法结构构建能做什么、不能做什么。《自然综述：神经科学》。https://doi.org/10.1038/s41583-022-00659-5

Keitel, A., Gross, J., & Kayser, C. (2018). 听觉皮层和运动皮层中与知觉相关的语音追踪反映了不同的语言特征。《公共科学图书馆·生物学》，

16(3)，e2004473。https://doi.org/10.1371/journal.pbio.2004473

Lo, C.-W., Tung, T.-Y., Ke, A. H., & Brennan, J. R. (2022). 层级结构而非词汇规则性调节语言理解过程中的低频神经同步。《语言神经生物学》，3(4), 538-555. https://doi.org/10.1162/nol_a_00077

Lu, L., Sheng, J., Liu, Z., & Gao, J. H.（2021）。频率标记脑磁图反应揭示的想象言语神经表征。《神经影像学》，229，117724。https://doi.org/10.1016/j.neuroimage.2021.117724

Makov, S., Sharon, O., Ding, N., Ben-Shachar, M., Nir, Y., & Zion Golumbic, E.（2017）。睡眠会扰乱高级言语解析，尽管基本听觉加工显著。《神经科学杂志》，37(32)，7772-7781。https://doi.org/10.1523/JNEUROSCI.0168-17.2017

Meng, Q., Hegner, Y. L., Giblin, I., McMahon, C., & Johnson, B. W.（2021）。清晰和退化言语中抽象语言结构的偏侧化大脑加工。《大脑皮层》，31(1)，591-602。https://doi.org/10.1093/cercor/bhaa245

Meyer, L. (2018). 言语加工与语言理解中的神经振荡：最新进展与新兴机制。《欧洲神经科学杂志》，48(7), 2609-2621. https://doi.org/10.1111/ejn.13748

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). 稳态视觉诱发电位在视觉研究中的应用：综述。《视觉杂志》，15(6), 4-4. https://doi.org/10.1167/15.6.4

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). 对拍子和节拍的神经元牵引进行标记。《神经科学杂志》，31(28), 10234-10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Oppenheim, A. V., Buck, J. R., & Schafer, R. W. (2001). 离散时间信号处理. 第2卷. 新泽西州上萨德尔河：Prentice Hall。

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D.（2003年）。人类听觉稳态反应：人类稳态听觉反应。《国际听力学杂志》，42(4)，177-219。

https://doi.org/10.3109/14992020309101316

Poor, H. V.（1998年）。信号检测与估计导论。Springer Science & Business Media。

Regan, D. (1977). 稳态诱发电位。《美国光学学会期刊》，67(11), 1475-1489. https://doi.org/10.1364/JOSA.67.001475

Ross, B., Borgmann, C., Draganova, R., Roberts, L., & Pantev, C. (2000). 一项关于人类听觉系统对幅度调制音的听觉稳态反应的高精度脑磁图研究。《美国声学学会期刊》，108, 679-691. https://doi.org/10.1121/1.429600

Wang, Y., Ding, N., Ahmar, N., Xiang, J., Poeppel, D., & Simon, J. Z. (2012). 人类听觉系统对时间调制率和频谱带宽的敏感性：脑磁图证据。《神经生理学杂志》，107(8), 2033-2041. https://doi.org/10.1152/jn.00310.2011

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). 音乐经验塑造人脑对语言音高模式的脑干编码。《自然-神经科学》，10(4), 420-422. https://doi.org/10.1038/nn1872

Yost, W. A. (1996). 迭代纹波噪声的音高。美国声学学会期刊，100(1), 511-518. https://doi.org/10.1121/1.415873

Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). 神经同步频域分析的解读：周期性、基频与谐波。《人类神经科学前沿》，10, 274. https://doi.org/10.3389/fnhum.2016.00274