Neural Entrainment to Rhythms of Imagined Syllables

Pan, C, Liu, H, Zheng, D & Chen, F

DOI 10.1109/EMBC48229.2022.9871767

Publisher: IEEE

2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

Neural Entrainment to Rhythms of Imagined Syllables

Changjie Pan, Haipeng Liu, Dingchang Zheng, Fei Chen, *Senior Member, IEEE*

Abstract— **Imagined speech based brain-computer interface (BCI) is of great interest due to its efficiency and user-friendliness for patients with speech impairment. The aim of this work was to study whether different rhythms of imagined syllables could elicit corresponding frequency components on EEG amplitude spectra. Seventeen participants were recruited to take part in the experiments, and performed a control task and four imagery tasks with the presence of periodic pure tones while their EEG signals were recorded. The four imagery tasks included imagining the syllable '/a/' every time, every two times, and every three times the periodic pure tones occurred, and imagined twice every three times the periodic pure tones occurred. The experimental results analyzed by Fourier transform indicated that neural entrainment to rhythmic speech imagery can be notably reflected on the EEG amplitude spectra.**

Clinical Relevance— **This work manifested that different rhythms of imagined syllables could be identified from EEG amplitude spectra, which may be beneficial to the development of imagined speech based BCIs.**

I. Introduction

Speech imagery, which is generally referred to inner speech, cover speech or imagined speech [1], is one of the speech modalities that human produce speech mentally and silently without any movement of vocal organs. Speech imagery is a common phenomenon that happens in our daily life, and it has been estimated that it takes up about more than one quarter of awaking time in lifetime [2].

In brain-computer interfaces (BCIs) domain, speech imagery based BCIs have been studied extensively in recent years [e.g., 3]. By classifying the recorded brain-related neurological signals when people imagine different categories of speech, speech based BCIs could generate a series of output commands. It has special as well as considerable merits to patients with speech impairments, strokes or locked-in syndromes. On one hand, those patients have difficulties in speech articulation, hence giving a clear and intelligible verbal instruction is troublesome for them. Speech imagery does not require any speech articulation and patients could connect with outside world conveniently just with their inner thoughts. One the other hand, human speech contains abundant information. Different vowels and consonants could form various syllables, and different syllables could form various words. Therefore, speech-based BCIs own the ability to generate much more output commands than traditional BCIs.

Based on the content of imagined speech, speech imagery based BCIs generally can be categorized into imagined syllable based BCIs [e.g., 4-5], imagined word based BCIs [e.g., 6-7] and imagined phrase based BCIs [e.g., 8]. However, studies about rhythmic imagined speech based BCIs are very limited. Deng et al. established a BCI by classifying the recorded EEG signals when the participants were asked to produce one of two syllables (i.e., '/ba/' and '/ku/') in one of three rhythms [9]. Watanabe et al. investigated the possibility of classifying three rhythms of imagined '/ba/' sounds by matching the speech envelope and the EEG signals at the speech imagery stage [10]. Though the classification results in the above study indicated the feasibility of rhythmic imagined speech based BCIs, the potential neural mechanism underlying rhythmic speech imagery was barely mentioned.

Recent studies showed that the rhythms of beat and meter imagery in music can be decoded from EEG amplitude spectra. Nozaradan et al. reported that binary (strong-weak) and ternary (strong-weak-weak) meter imagery that imposed on beat could elicit corresponding frequency components on EEG spectra at the frequency of subharmonics of beat in complex sound [11]. Celma-Miralles et al. also demonstrated similar results with the presence of periodic attenuated pure tones. Besides, they further extended the knowledge regarding neural entrainment to meter imagery from auditory domain to visual domain [12].

The rhymical imagined speech is similar to meter imagery in musical perception to some extent. Consequently, it is hypothesized that different rhythms of imagined syllables will elicit corresponding frequency components in EEG amplitude spectra. The findings of this work will provide important evidence and insight for the design of classification models in rhymical speech imagery BCIs.

II. Methods

A. Participants

Seventeen Mandarin-Chinese participants (5 females and 12 males, and ranging from 19 to 26 years old) participated in the experiment. All participants had normal hearing abilities and had no history of neurological or psychological disorders. All participants had signed the informed consents and the Research Ethics Committee of Southern University of Science and Technology had reviewed and approved this study.

B. Stimulus

While participants were performing different tasks, the auditory stimuli were presented to provide the reference pace. The auditory stimuli were an isochronous sequence of 15 1000-Hz pure tones, which lasted for about 12 seconds in total.

This work was supported by Shenzhen Key Laboratory of Robotics Perception and Intelligence (ZDSYS20200810171800001), the National Natural Science Foundation of China (61971212), and Shenzhen Sustainable Support Program for High-level University (20200925154002001).

Changjie Pan and Fei Chen are with Shenzhen Key Laboratory of Robotics Perception and Intelligence, and the Department of Electronic and Electrical Engineering, Southern University of Science and Technology, Shenzhen 518055, China (Phone: +86-755-88018554, Email: fchen@sustech.edu.cn).

Haipeng Liu and Dingchang Zheng are with Research Centre for Intelligent Healthcare, Coventry University, Coventry, CV1 5FB, United Kingdom.

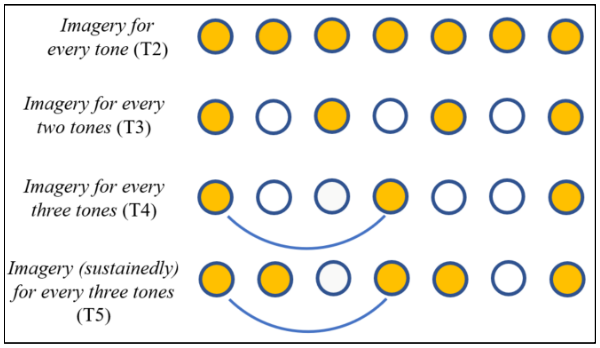

Figure 1. Illustration of the imagined rhythms in the four imagery tasks (T2, T3, T4, and T5) in this work. Each blue ring indicates the moment when a pure tone occurs, and the yellow filled circles indicate the moment when participants imagine the sound '/a/'.

The pure tones were presented at a frequency of 1.2 Hz, i.e., the interval for two adjacent pure tones was 833.33 ms, and each pure tone lasted for 100 ms. The auditory stimuli were presented to the participants binaurally at a comfortable intensity level [~60 dB(A) sound pressure level (SPL) over a noise floor of <40 dB(A) SPL] through an earphone.

C. Procedure

In the experiment, each participants performed five different tasks: 1) the *control* task (T1). The participants were asked to listen to the auditory stimuli attentively and think about nothing. 2) the *Imagery for every tone* task (T2). The participants were required to imagine the pronunciation of the syllable '/a/' without any involvement of vocalization immediately when they hear a pure tone. The frequency of speech imagery was 1.2 Hz. 3) the *Imagery for every two tones* task (T3). The participants imagined the syllable '/a/' every two times they heard the pure tone. The frequency of speech imagery was 0.6 Hz. 4) the *Imagery for every three tones* task (T4). The participants imagined the syllable '/a/' every three times they heard the pure tone. The frequency of speech imagery was 0.4 Hz. 5) the *Imagery (sustainedly) for every three tones* task (T5). The participants imagined the syllable '/a/' every three times they heard the pure tone, and they also imagined the syllable '/a/' at the moment the next pure tone occurred. Before the experiment started, all participants were trained to control their imagination duration to be the interval of two adjacent pure tones as the same as possible in the training block. The illustration of the imagined rhythms in the four imagery tasks (i.e., T2, T3, T4, and T5) was shown in Fig. 1.

The experiment consisted of 5 blocks, and each block contained 30 trials. The duration of each block was about 12 minutes. Each task consisted of 30 trials, which was distributed randomly across all blocks. At the beginning of each trial, the participants were instructed to look at the screen for 4 seconds, during which the screen displayed the task name for the current trial. Then the screen was blank, after 2 seconds, the auditory stimuli were presented to the participants binaurally and the participants were instructed to perform the corresponding task till all auditory stimuli ended. Afterwards, the participants were asked to report whether the task was performed successfully through a button board. Once the button was pressed, the experiment would move onto the next trial.

Table 1. Results of one-tailed pairwise Wilcoxon signed-tank tests for amplitude of target frequencies for four imagery tasks over control task (*$p < 0.05$, **$p < 0.01$).

Imagery task	Target frequency (Hz)	*z*-value	*p*-value
Imagery for every tone	0.4	-1.874	0.030*
(T2 over T1)	0.6	-0.568	0.285
	0.8	-0.852	0.197
	1.2	-1.761	0.039*
Imagery for every two tones	0.4	-1.874	0.030*
(T3 over T1)	**0.6**	-0.568	0.285
	0.8	-0.852	0.197
	1.2	-1.761	0.039*
Imagery for every three tones	**0.4**	-3.124	0.001**
(T4 over T1)	0.6	-0.568	0.285
	0.8	-1.761	0.039*
	1.2	-2.215	0.013*
Imagery (sustainedly) for every three tones	**0.4**	-2.101	0.018*
(T5 over T1)	0.6	-1.306	0.096
	0.8	-2.385	0.008**
	1.2	-1.42	0.078

Table 2. Results of two-tailed pairwise Wilcoxon signed-tank tests for amplitude of target frequencies for *every three tones* task and *imagery (sustainedly) for every three tones* task (T4 and T5) (*$p < 0.05$, **$p < 0.01$).

Target frequency (Hz)	*z*-value	*p*-value	Amplitude comparation
0.4	-2.442	0.015*	T4>T5
0.6	-1.250	0.211	n.s.
0.8	-0.682	0.496	n.s.
1.2	-1.363	0.173	n.s.

D. EEG data recording and pre-processing

The participants were seated comfortably with an EEG cap on the head in an acoustically and electrically shielded chamber during the experiment. They were required to minimize their movement to avoid possible artifact contamination in the recorded EEG data. The EEG elastic cap (NeuroScan Inc.) had 64 electrodes, and was placed at specific positions following the extended international 10-20 system. The ground electrode was attached to the forehead and the reference electrode was placed at the top of the nose. The sampling rate for EEG data was 1000 Hz and the impedance between any recording electrode and the reference electrode was maintained below 5 kΩ.

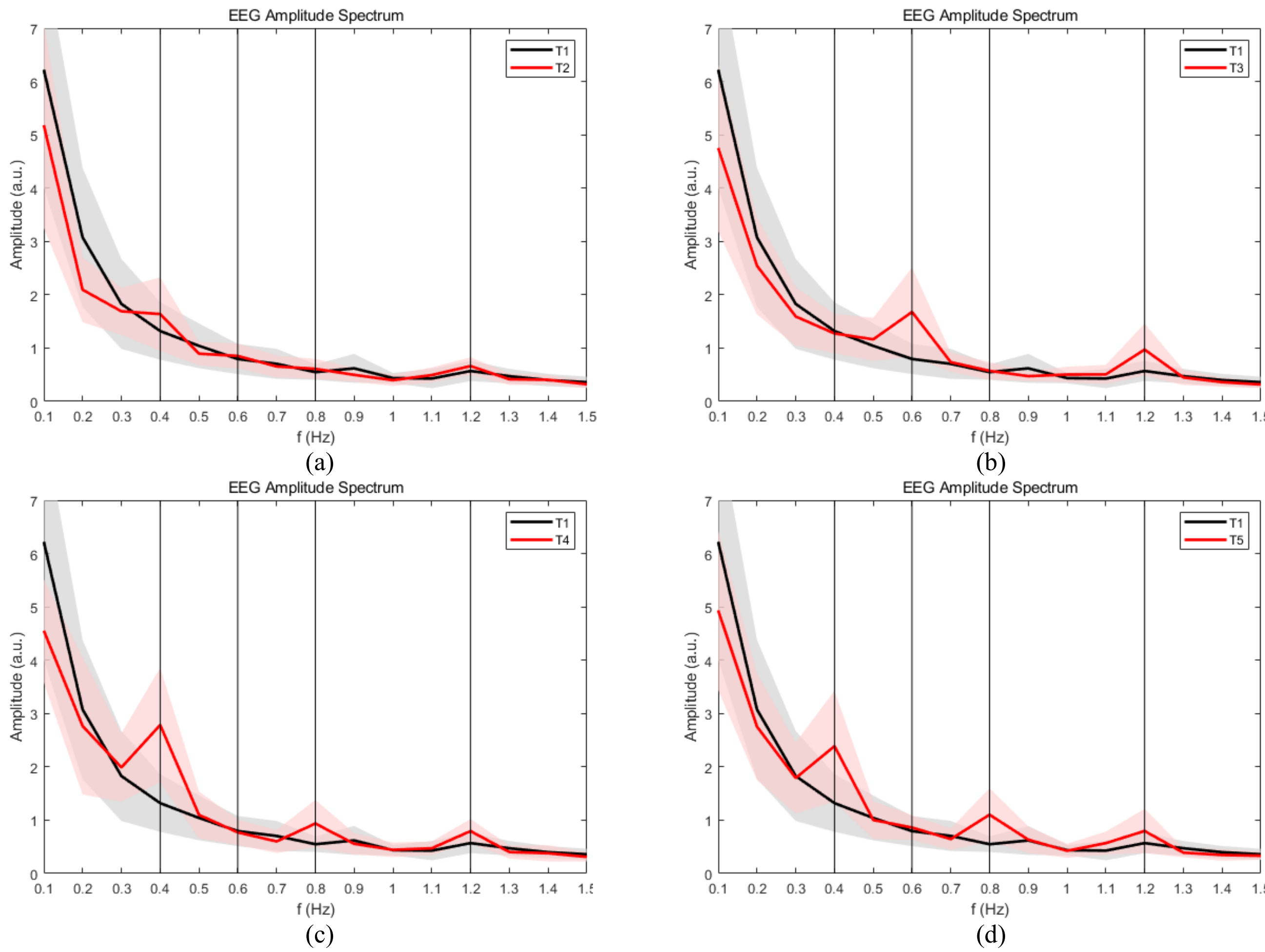

Figure 2. The average amplitude spectra across all participants of the *Imagery for every tone* task (T2), the *Imagery for every two tones* task (T3), the *Imagery for every three tones* task (T4), and the *Imagery (sustainedly) for every three tones* task (T5) in comparison with the control task (T1) are shown in (a), (b), (c), and (d), respectively. The black line represents the averaged amplitude spectra for the *control* task, and the red lines represent the averaged amplitude spectra for the four imagery tasks. The shaded areas represent the 95% confidence bands of mean power energies.

EEGLAB toolbox in MATLAB was applied to do the preprocessing on the raw EEG data. The EEG data were first re-referenced using electrodes at contralateral mastoid, and then band-pass filtered between 0.1 Hz and 10 Hz. The artifacts (e.g., electrocardiographic activities, eye blinks, horizontal eye movements, etc.) in EEG data were removed by Independent Component Analysis (ICA). After the artifact removal, the epochs in the interval that the auditory stimulus was presented were extracted between 1-s pre-stimulus and 10-s post-stimulus, and then corrected with the baseline of 1-s pre-stimulus. One female and one male's EEG data were excluded from subsequent analysis due to the poor data quality. The heavily-contaminated epochs and those epochs that the tasks were not successfully performed were also excluded from EEG spectra analysis.

E. EEG spectra analysis

For each task and each participant, the 10-s pre-processed EEG data were averaged both across all epochs and across all electrodes. Then the averaged EEG data were transformed into frequency domain by applying discrete Fourier transform. The frequency resolution for the obtained EEG amplitude spectra was set to be 0.1 Hz.

III. RESULTS

The averaged amplitude spectra across all participants from frequency 0.1 Hz to 1.5 Hz were presented in Fig. 2. For each imagery task (i.e., T2 to T5), one-tailed pairwise Wilcoxon signed-rand tests were conducted to determine whether the amplitudes of imagery tasks were significantly larger than that of control task at the target frequencies of 0.4 Hz, 0.6 Hz, 0.8 Hz, and 1.2 Hz. Target frequency of 1.2 Hz corresponded to the frequency of auditory stimulus, as well as the frequency of the *Imagery for every tone* task (T2). Target frequency of 0.6 Hz corresponded to the frequency of the *Imagery for every two tones* task (T3). Target frequency of 0.4 Hz corresponded to the frequency of the *Imagery for every three tones* task (T4) and the *Imagery (sustainedly) for every three tones* task (T5). The frequency of 0.8 Hz, i.e., the second harmonic of 0.4 Hz, was also been analyzed since this frequency component was observed during the ternary meter imagery as reported in [11-12].

The results of one-tailed pairwise Wilcoxon signed-tank tests for amplitude of target frequencies for imagery tasks over control task are demonstrated in Table 1. In the *Imagery for every tone* task, the EEG spectral amplitudes at target frequencies of 0.4 Hz and 1.2 Hz are significantly higher than those in the *control* task (0.4 Hz: $z = -1.874, p = 0.030$; 1.2 Hz: $z = -1.761, p = 0.039$). In the *Imagery for every two tones* task, the EEG spectral amplitudes of 0.6 Hz and 1.2 Hz are significantly higher than those in the *control* task (0.6 Hz: $z = -2.158$, $p = 0.015$; 1.2 Hz: $z = -1.988$, $p = 0.023$). In the *Imagery for every three tones* task, the EEG spectral amplitudes of 0.4 Hz, 0.8 Hz and 1.2 Hz are higher than those in the *control* task (0.4 Hz: $z = -3.124, p = 0.001$; 0.8 Hz: $z = -1.761, p = 0.039$; 1.2 Hz: $z = -2.215, p = 0.013$). In the *Imagery (sustainedly) for every three tones* task, the EEG spectral amplitudes of 0.4 Hz and 0.8 Hz are higher than those in the *control* task (0.4 Hz: $z = -2.101, p = 0.018$; 0.8 Hz: $z = -2.385$, $p = 0.008$).

Moreover, two-tailed pairwise Wilcoxon signed-rank test was applied to compare the EEG spectral amplitudes of the *Imagery for every three tones* task and the *Imagery (sustainedly) for every three tones* task (T4 and T5). The results are shown in Table 2. It is shown that the EEG spectral amplitude in the *Imagery for every three tones* task is significantly greater than that in the *Imagery (sustainedly) for every three tones* task at the target frequency of 1.2 Hz ($z = -2.442, p = 0.015$). The differences for EEG spectral amplitude at the target frequency of 0.4 Hz, 0.6 Hz and 0.8 Hz are not significant.

IV. Discussions and Conclusion

As indicated in Fig. 2 and Table 1, relative to the *control* task, the EEG spectral amplitudes at frequency of 1.2 Hz were observed (and statistically compared) to be significantly different by the Wilcoxon signed-rank test in most of imagery tasks (except for the *Imagery (sustainedly) for every three tones* task). These results demonstrate that imagery tasks usually tend to enhance the EEG spectral amplitudes at the frequency of external auditory stimulus (i.e., 1.2 Hz in this work). This is consistent with findings in previous studies [e.g., 11-12], which demonstrated that peaks in EEG amplitude spectra at the frequency of beat could be observed in all imagery tasks.

In the *Imagery for every two tones* task, the significant peak occurred at the frequency of the binary subharmonics of the auditory stimuli, i.e., 0.6 Hz as presented in Fig. 2 (b) and Table 1. In the *Imagery for every three tones* task, significant frequency component was observed at the frequency of the ternary subharmonics of the auditory stimulus (i.e., 0.4 Hz), and the second harmonics of 0.4 Hz (i.e., 0.8 Hz), as indicated in Fig. 2 (c) and Table 1. These results manifested that the neural entrainment to rhythms of imagined syllables in binary or ternary paces could be reflected on the EEG spectra, which were consistent with the neurological findings in meter imagery [11-12].

The *Imagery for every three tones* task (T4) and the *Imagery for every three tones* task (T5) both corresponded to the rhythm of frequency of 0.4 Hz, and their EEG amplitude spectra were similar – frequency components of 0.4 Hz and 0.8 Hz were observed – as indicated in Fig. 2 (c) and (d). In order to investigate whether these two imagery tasks owned distinguished spectral characteristics, two-tailed pairwise Wilcoxon signed-rank test was applied as shown in Table 2. The results indicated that the sole difference between the two spectra was that the EEG amplitude in the *Imagery for every three tones* task at the target frequency of 0.4 Hz was higher than that in the *Imagery (sustainedly) for every three tones* task. This suggests that imagery for one time requires more attention in brain than imagery for two times at the same rhythm of speech imagery, and hence the spectral energy is enhanced at the frequency of the corresponding rhythm in EEG spectra, which warrants further investigation in future.

In conclusion, this study manifested that the different rhythms of imagined syllables in the presence of external auditory stimuli could elicit corresponding frequency components in EEG spectra. Furthermore, the EEG amplitude spectra of four different imagery tasks differed distinctly from each other, not only on frequency ranges but also on spectral amplitudes. The findings of this study indicated that the classification method in frequency domain may be suitable for the design of rhythm-based imagined speech BCIs.

References

[1] M. Perrone-Bertolotti, L. Rapin, J.P. Lachaux, M. Baciu, and H. Lœvenbruck, "What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring," *Behav. Brain Res.*, vol. 261, pp. 220–239, 2014.

[2] C.L. Heavey, R.T. Hurlburt, "The phenomena of inner experience," *Conscious Cogn.*, vol. 17, no. 3, pp. 798–810, 2008.

[3] B. Denby, T. Schultz, K. Honda, T. Hueber, J. Gilbert, and J. Brumberg, "Silent speech interfaces," *Speech Commun.*, vol. 52, no. 4, pp. 270–287, 2010.

[4] C.S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural networks*, vol. 22, no. 9, pp. 1334–1339, 2009.

[5] B. Min, J. Kim, H.J. Park, and B. Lee, "Vowel imagery decoding toward silent speech BCI using extreme learning machine with electroencephalogram," *BioMed Research International*, vol. 2016, pp. 2618265, 2016.

[6] A.R. Sereshkeh, R. Trott, A. Bricout, and T. Chau, "Online EEG classification of covert speech for brain-computer interfacing," *International Journal of Neural Systems*, vol. 27, no. 8, pp. 1750033, 2017.

[7] C.H. Nguyen, G.K. Karavas, and P. Artemiadis, "Inferring imagined speech using EEG signals: a new approach using Riemannian manifold features," *Journal of Neural Engineering*, vol. 15, no. 1, pp. 016002, 2018.

[8] D. Dash D, P. Ferrari, and J. Wang, "Decoding imagined and spoken phrases from non-invasive neural(MEG) signals," *Frontiers in Neuroinformatics*, vol. 14, pp. 290, 2020.

[9] S. Deng, R. Srinivasan, T. Lappas, and M. D'Zmura, "EEG classification of imagined syllable rhythm using Hilbert spectrum methods," *Journal of neural engineering*, vol. 7, no. 4, pp. 046006, 2010.

[10] H. Watanabe, H. Tanaka, S. Sakti, and S. Nakamura, "Synchronization between overt speech envelope and EEG oscillations during imagined speech," *Neuroscience research*, vol. 153, pp. 48–55, 2020.

[11] S. Nozaradan, I. Peretz, M. Missal, and A. Mouraux, "Tagging the neuronal entrainment to beat and meter," *The Journal of neuroscience*, no. 31, vol. 28, pp. 10234–10240, 2011.

[12] A. elma-Miralles, R. F. de Menezes, and J. M. Toro, "Look at the beat, feel the meter: top-down effects of meter induction on auditory and visual modalities," *Frontiers in human neuroscience*, vol. 10, pp. 108, 2016.