

# Optimizing Auditory Input for Foreign Language Learners Through A Verbotonal-Based Dichotic Listening Approach

**Xirui Cai**

Suranaree University of Technology <https://orcid.org/0000-0003-1158-5895>

**Andrew Lian**

Suranaree University of Technology

**Nattaya Puakpong**

Suranaree University of Technology

**Yaoping Shi**

Kunming Medical University First Affiliated Hospital

**Haoqiang Chen**

Yunnan Provincial People's Hospital: First People's Hospital of Yunnan

**Yizhen Zeng**

Kunming Medical University First Affiliated Hospital

**Jibing Ou**

Kunming Medical University Second Hospital

**Wei Zheng**

South-Central University for Nationalities

**Yin Mo** (✉ [likecc8711@msn.com](mailto:likecc8711@msn.com))

The First Affiliated Hospital of Kunming Medical University

---

## Original article

**Keywords:** verbotonalism, dichotic approach, ERP, fMRI, ESL learners

**Posted Date:** May 5th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-449878/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Asian-Pacific Journal of Second and Foreign Language Education on September 1st, 2021. See the published version at

<https://doi.org/10.1186/s40862-021-00119-0>.

# Abstract

The quality of the physical language signals to which learners are exposed and which result in neurobiological activity leading to perception should be prioritized as it constitutes a variable that is rarely, if ever, considered in language learning. The current study identifies an optimal audio language input signal for Chinese EFL/ESL learners generated by modifying the physical features of language-bearing audio signals. This is achieved by applying the principles of verbotonalism in a dichotic listening context. Low-pass filtered (320 Hz cut-off) and unfiltered speech signals in four different configurations were selectively directed to each hemisphere of the brain through the contralateral ear. The four types of auditory stimuli were: low-pass filtered stimuli in both ears (FL-FR), filtered stimuli in the left ear and unfiltered stimuli in the right ear (FL-R), unfiltered stimuli in the left ear and filtered in the right ear (L-FR), and unfiltered stimuli in both ears (NL-NR). Temporal and spatial neural signatures for the processing of the signals were detected in a combined Event-Related Potential (ERP) and functional Magnetic Resonance Imaging (fMRI) experiment. Results showed that the FL-R configuration provided optimal auditory language input by actively exploiting left-hemispheric dominance for language processing and right-hemispheric dominance for melodic processing, i.e., each hemisphere was fed the signals that it should be best equipped to process and it actually did so effectively. In addition, the L-FR configuration was identified as entirely non-optimal for language learners. Other outcomes included significant load reduction through exposure to FL-FR signals as well as the confirmation that non-language signals were recognized by the brain and did not trigger any language processing. These various outcomes will necessarily entail further research.

## 1. Background

Language learning is essentially an individual meaning-making process, during which learners' perceptions of language signals are constructed into meanings (Lian, 2004). Thus, the goal of foreign language instruction is to enable students to make better sense of foreign language signals and, as a result, to structure/re-structure their perceptions of the language being learned to facilitate language learning. It is common for foreign language teachers to put much emphasis on higher-order cognitive skills, such as applying rules, communicating in oral and written forms, translating to or from the target language, or even reading and thinking critically in a foreign language (Galotti, 2017; Levine, 2009). These seem to place heavy demands on foreign language learners with intermediate or lower levels of language proficiency. However, the importance of raising students' awareness of foreign language signals has always been neglected. In fact, the biological basis, especially the brain, is the primary perceiver of the language signals during the learning process. Thus, perceptions of physical language signals resulting from the neurobiological activity should be prioritized at least as much as activities requiring higher-level cognitive skills especially for foreign language learners.

As a key factor to raise students' awareness of a (foreign) language, an optimal language input that is best suited for neural processing will contribute significantly to making effective learning happen, i.e., it would maximize the neuroplasticity. For language input in second language acquisition, Krashen (1982,

1985, 2003) initiated the input hypothesis that the optimal input for a second language learner is “comprehensible input”. However, as McLaughlin (1987) criticizes, the concept of “comprehensible input” is not precisely defined, and the ambiguity of “ $i+1$ ” makes the hypothesis untestable. Further, the physical features of the optimal language input that may assist language learning are not mentioned in the hypothesis. It is exclusively a linguistic hypothesis.

In order to provide language learners and hearing-impaired subjects with optimal language input, Guberina proposed the Verbotonal approach (Guberina & Asp, 1981, 2013) that considers language/speech development as a meaning-making process (Lian, 1980, 2004, 2011, 2014). In addition, the way a speaker produces speech reflects how he/she perceives speech (Guberina & Asp, 1981, 2013). Thus, changing learners’ perceptions of speech would give rise to changes in their speech production. In other words, the correction of the speakers’ speech would be based upon the correction of speech perception. The Verbotonal approach emphasizes that prosodic information (intonation and rhythm), contained in low frequencies, conveys meanings and changes learners’ perception and production of speech (Guberina & Asp, 1981, 2013; Kim & Asp, 2002). The reason is that the cochlea and the vestibular organ develop from the feeling of speech prosody in low frequencies during fetal life, which is the basis of the proprioceptive memory and auditory-memory development (Asp, 2006). Thus, our ears are sensitive to the changes of pitch, rhythm, and intonation via low frequencies. Low-frequency signals are derived from a low-pass filter at a cutoff frequency of about 320 Hz, in which the fundamental frequency ( $F_0$ ) and prosodic features (stress, rhythm, loudness, and intonation) are maintained but the frequencies (above 320 Hz) that make words identifiable are removed (Lian & Sussex, 2018). From the clinical practice of speech-language pathology and the practice of foreign language learning, the Verbotonal approach with low-pass filtered signals is identified as effective for rehabilitating hearing-impaired children (e.g., Asp, Kim, & Davis, 2003; Jurjević-Grkinić, Munivrana, & Mijić, 2015) and for teaching ESL learners. ESL learners’ pronunciation was improved significantly after the Verbotonal approach was implemented, including English speaking skills in general (He, 2014; He et al., 2015; Yang, 2016; Yang et al., 2017), pronunciation correction (Wen et al., 2020), as well as phonological working memory (Yang, 2016; Yang et al., 2017). Previous studies provide empirical evidence that the Verbotonal approach with low-pass filtered signals offers an optimal model for language perception and production, which may re-structure learners’ perceptions of speech and ultimately lead to maximal neuroplasticity for language learning. However, research on the neural processing mechanisms of the low-pass filtered signals is still scarce. Thus, the current study aims to unveil a Chinese ESL learner’s brain activity in response to low-pass filtered and unfiltered signals in a combined ERP (event-related potential) and fMRI (functional magnetic resonance imaging) investigation.

As mentioned above, an optimal language input signal ought to be the signal that is best suited for the brain to process. This principle will form the basis for the research reported here. Specifically, our assumption is based on ear advantage for linguistic and melodic signal processing, in which the underlying neural mechanisms suggest that the left hemisphere that connects primarily to the right ear is dominant in the language signal processing (e.g., Tervaniemi & Hugdahl, 2003; Vigneau et al., 2006), and

the right hemisphere that links to the left ear is predominant in the melodic and speech intonation processing (e.g., Meyer et al., 2002; Sammler et al., 2015). A dichotic listening approach is then adopted to selectively direct the low-pass filtered language signals with only prosodic features of speech through the left ear to the right brain, and the unfiltered signals with linguistic information are selectively directed through the right ear to the left brain. It is assumed, from a verbotonal perspective that, in this configuration, optimal language signals are sent to each hemisphere of the brain. The detailed descriptions of the sentences used as auditory stimuli and the configurations of the stimuli will be illustrated in 2.2.

ERP and fMRI technologies rely on noninvasive techniques widely used to track brain responses to language signals. ERP measures electrophysiological responses to cognitive-related events, presented by brain wave activities and the components elicited by the experimental tasks or events (Antonenko et al., 2014; Luck, 2014). A negative wave peaking at 400 ms is known as N400, which is usually elicited by the semantic violation in a sentence; and a positive wave peaking at 600 ms is identified as P600 that is elicited by the syntactic violation (Daltrozzo & Conway, 2014; Kutas & Federmeier, 2011; Morgan-Short & Tanner, 2014; Swaab et al., 2012). These two ERP components are language-related regarding semantic and syntactic manipulations and will be used extensively in this study. These elicited components can be analyzed by their amplitude, latency, and distribution to understand the cognitive processes related to the experimental tasks or events (Antonenko et al., 2014). fMRI detects changes in blood oxygenation in response to neural activation with high spatial-resolution images (Kuhl & Rivera-Gaxiola, 2008; Logothetis, 2012). The activated brain regions induced by a specific task or activity are determined by the blood oxygenation level dependent (BOLD) signals (Indefrey, 2012; Poldrack, 2018). Functional localization of processing the low-pass filtered and unfiltered language signals would be unveiled in the fMRI experiment. Thus, the current case study employs a combined ERP and fMRI approach that aims to provide temporal and spatial neural signatures for the processing of the low-pass filtered and unfiltered English sentences under dichotic listening conditions in a Chinese ESL student. As a result, an optimal language input for Chinese ESL learners is expected to be identified within the Verbotonal-based dichotic approach.

## 2. Methods And Materials

### 2.1 Participant

A male graduate student aged 27 years took part in the current study. He was strongly right-handed with a laterality index = 100 according to the Edinburgh Handedness Inventory (Oldfield, 1971), had normal hearing and vision, no neurological or psychiatric disease, and no contraindication for MRI scanning. Mandarin Chinese was the participant's first language, and he had been learning English as a foreign language for 15 years. As for English language proficiency, he had passed the standardized English proficiency test for college students in China (the College English Test, CET-Band 4, the medium level of the test), and self-evaluation of English proficiency was intermediate level. The participant gave written informed consent before participating in the study, which was approved by the local Ethics Committee.

## 2.2 Auditory language stimuli

The listening materials, presented in this experiment, were complete English sentences extracted from the Cambridge English qualifying exam – Cambridge English Preliminary (PET) with B1/intermediate difficulty level. Since the participant's actual English listening proficiency was basically consistent with the difficulty level of PET listening materials, listening materials included in *Objective PET Teacher's Book (4th Edition)* (Hashemi & Thomas, 2013) would impose a mental workload consistent with his level of knowledge. He should be able to handle the load with neither great difficulty nor great ease. The sentences were pronounced by two native speakers (one male and one female) at a rate of 200 words per minute approximately. Sentences used as auditory stimuli in the current study are listed in Table 1. The mean number of words in the sentences was 9.33 (SD 0.82) and the mean duration of the signals was 2638.33 (SD 147.30) ms. All auditory stimuli were designed to be 3000 ms, during which sentence signals lasted  $2638.33 \pm 147.30$  ms, followed by the silence of  $361.67 \pm 147.30$  ms.

**Table 1.** Sentences used as auditory language stimuli

Sentences	Words (N)	Duration (ms)
I was wondering if you could get me a newspaper.	10	2450
I preferred baseball when I was at school.	8	2560
I'm happy just doing the things I can do.	9	2600
We went on a trip to swim with dolphins.	9	2600
There's a new computer game shop which is just opened.	10	2760
My parents want me to study to be a doctor.	10	2860

The auditory stimuli consisted of filtered and unfiltered sounds in both ears, organized into four configurations of stimuli: (a) filtered stimuli in both channels (FL-FR); (b) filtered stimuli in the left channel and unfiltered stimuli in the right channel (FL-R); (c) unfiltered stimuli in the left channel and filtered in the right channel (L-FR); (d) unfiltered stimuli in both channels (NL-NR). All auditory stimuli were edited by using Adobe Audition (Version 11.1.0; <https://adobe.com/products/audition>) at a 44.1 kHz sampling rate in a 32-bit stereo audio track. Low-pass filtering was obtained by setting a cut-off frequency of 320 Hz as per standard practice previous verbotonal experiments. 320 Hz low-pass filtering means that frequencies above 320 Hz were removed from the sound signals and frequencies below 320 Hz were maintained. According to Meyer et al. (2002), amplitudes of the unfiltered stimuli were normalized to 70%, and amplitudes of the low-pass filtered signals were normalized to 85% to ensure equal intensity as filtered sounds were limited in bandwidth and had lost energy as a result of the filtering. Figure 1 shows the sentence "I preferred baseball when I was at school." with FL-R configuration.

## 2.3 ERP experiment

### 2.3.1 ERP experimental design

The Oddball paradigm with three types of stimuli was adopted in the current study, which contained frequent non-target standard stimuli, less frequent target deviant stimuli, and rare non-target novel stimuli. As the research focus of the current study was the neural processing of the low-pass filtered auditory language stimuli under dichotic conditions, the deviant stimuli were designed to be FL-R, L-FR, and FL-FR signals in three separate runs of ERP recordings. In addition, in three runs of the experiment, standard stimuli consisted of NL-NR signals and novel stimuli consisted of non-linguistic environmental sounds (running water, train noises, and birds singing). Three runs of ERP recording were performed to investigate three configurations of the low-pass filtered stimuli (i.e., FL-R, L-FR, and FL-FR signals) respectively, which meant that one configuration of stimuli was examined in one run. To better spot the differences in auditory language processing mechanisms for different configurations of signals, the listening materials were the same but the only difference was the configurations of low-pass filtering and unfiltering in the left or/and right channel(s) in three runs. As a result, four configurations of stimuli (FL-FR, FL-R, L-FR, and NL-NR) in three separate runs of the ERP experiment would reveal the auditory language processing mechanisms for dichotic listening to filtered and unfiltered signals. In total, 210 auditory stimuli were presented in each run, including 150 standard NL-NR stimuli (with a frequency of occurrence of 71.4%), 50 deviant stimuli with one of the FL-R, L-FR, and FL-FR configurations (23.8%), and 10 novel stimuli of environmental sounds (4.8%). Stimulus presentation was controlled by E-Prime 3.0 (<https://pstnet.com/products/e-prime>), and all auditory stimuli were presented randomly and continuously to the participant. Each run took 630 seconds and the whole ERP experiment took around 40 minutes.

### 2.3.2 ERP data acquisition

The participant was fitted with an electrode cap with tin electrodes based on the International 10-20 System (Klem, Lüders, Jasper, & Elger, 1999). Data was acquired from the sixteen electrode sites of Fp1, Fp2, F3, Fz, F4, T3, C3, Cz, C4, T4, P3, Pz, P4, O1, Oz, and O2 (as shown in Figure 3). During the experiment, auditory stimuli were presented via stereo headphones (Sennheiser HD 435), and the participant was asked to listen to the signals and close his eyes to avoid blinks. During data recording, electrode impedance was maintained below 20 k $\Omega$ . A 30-second recording of resting-state EEG data was used as a control (baseline) signal. As each auditory signal was within 3000 ms, the duration of data collection for each stimulus was 3000 ms. EEGLAB (Version 15.0.0b; <https://sccn.ucsd.edu/eeglab>) was used to analyze the EEG data. A band-pass filter of 0.1-70.0 Hz was adopted to filter the off-line data, and artifacts above 100  $\mu$ V were rejected.

### 2.3.3 ERP data analysis

The ERP components of N400 and P600 are language-related ERP components regarding semantic and syntactic manipulations (Daltrozzo & Conway, 2014; Kutas & Federmeier, 2011; Morgan-Short & Tanner, 2014; Swaab et al., 2012). In addition, the N400 and P600 effects presented typical centroparietal scalp distribution, which were maximal at midline centroparietal sites (e.g., Brouwer et al., 2012; Brouwer & Hoeks, 2013; Swaab et al., 2012; van Herten et al., 2005). Thus, the measures of amplitudes, latencies, and scalp topographies of N400 and P600 components within 1000 ms after the stimulus onset at the midline sites (Fz, Cz, and Pz) were taken for main analyses in the current study.

For amplitude analysis, increased mental processing load led to increases in the amplitudes of the N400 and P600 components. The reason for this was that the larger N400 amplitude was elicited by semantically less predictable and difficult information, and the larger P600 amplitude was evoked by syntactically difficult or less-preferred sentences (Kutas & Federmeier, 2011; Swaab et al., 2012). Thus, the amplitudes of N400 and P600 reflected mental workload for sentential processing in the current study. As to latency analysis, the latencies of N400 and P600 components indicated response time for processing the auditory stimuli, which reflected the stimulus evaluation process and the relative timing of the response to the signals (Swaab et al., 2012). Scalp topographies of N400 and P600 components demonstrated distributions of the electrical activity over the scalp that were elicited by the auditory stimuli. As a result, the amplitudes, latencies, and scalp topographies of the N400 and P600 components would reveal the mental processing workload, response time, and distributions of electrical activity while manipulating semantic and syntactic processing of the low-pass filtered and unfiltered signals under dichotic listening conditions.

## 2.4 fMRI experiment

### 2.4.1 fMRI experimental design

A block design was adopted in the fMRI experiment, which was innately suited for the detection of the brain regions activated by particular tasks/stimuli compared to other paradigms (Donaldson, 2004; Petersen & Dubis, 2012). According to the four configurations of the auditory language signals (i.e., FL-FR, FL-R, L-FR, and NL-NR), there were four runs of fMRI scanning. Each run of fMRI scanning examined one configuration of auditory signals. The block design in the fMRI experiment contained rest and stimulus blocks in each run, in which one rest/stimulus block lasted 12 seconds. A twelve-block design, i.e., six rest blocks and six stimulus blocks, were used as the paradigm and presented continuously and alternately, which took 144 seconds in each run of fMRI scanning for each configuration of auditory language signals (as shown in Figure 4A).

In each stimulus block, there was one sentence which was repeated four times as four auditory trials. Each auditory trial was within 3 seconds (as illustrated in Table 1) and one auditory trial was formulated to be 3 seconds, thus, one stimulus block contained four auditory trials and lasted 12 seconds. Six sentences were included in the six stimulus blocks respectively. Between stimulus blocks, there was a 12-

second silence designed to be a rest block as the baseline. Therefore, six stimulus blocks contained six sentences, totaling 24 auditory trials, and 12-second rest blocks were intervals between the stimulus blocks. The rest-stimulus block design for the fMRI experiment is illustrated in Figure 4B.

## 2.4.2 fMRI data acquisition

All images were acquired by using a General Electric MR750w 3.0T MRI scanner (GE, USA). The participant lay in the scanner with his head position secured with foam padding, wearing MRI-compatible pneumatic in-ear headphones. A compression alarm ball was placed in the participant's dominant right hand. The head coil was positioned over the participant's head to avoid motion during scanning. During scanning, there was no task for the participants. The participant was asked to be relaxed and listen to each signal.

Gradient echo (GRE) localizer images were acquired to determine the placement of the functional slices. For the functional images, a susceptibility weighted single-shot echo planar imaging (EPI) method with blood oxygenation level-dependency (BOLD) was used with the following scan parameters: TR = 2000 ms, TE = 30 ms, flip angle = 90°, matrix size = 64 × 64, field of view = 224 × 224 mm, slice thickness = 3.0 mm, gap = 1 mm, number of slices = 36. These parameters led to a 3.5 × 3.5 × 3.0 mm voxel size. By using an interleaved bottom-to-top sequence, 72 whole-brain volumes were acquired for each run. In order to get a high resolution, T1 weighted 3D images were obtained via a 3D magnetization-prepared GRE sequence with the following parameters: TR = 8.5 ms, TE = 3.2 ms, flip angle = 12°, matrix size = 256 × 256, field of view = 256 × 256 mm, slice thickness = 1 mm, gap = 0 mm, number of slices = 148, resulting voxel size = 1 × 1 × 1 mm. The acquisition of the anatomical scan and four runs of the fMRI scan took approximately 30 minutes.

## 2.4.3 fMRI data analysis

The fMRI data was analyzed by using the Statistical Parametric Mapping software (SPM12; <https://www.fil.ion.ucl.ac.uk/spm/software/spm12>) (Penny et al., 2011) developed by Friston et al. (1995). The functional images in the DICOM (Digital Imaging and Communications in Medicine) format were first converted to SPM-compatible image rolls via the DICOM Import option. Data preprocessing included slice timing correction, realignment for estimation and reslicing (the rejection threshold for excessive motion was a three-dimensional motion exceeding 1.5 mm or a three-dimensional rotation exceeding 10 degrees), normalization, and smoothing. Data that met the requirements were analyzed statistically.

For statistical analysis, the probability threshold was set at  $p < 0.001$  (uncorrected), and the pattern of BOLD-fMRI activation at a 10-voxel cluster threshold. The functional activation images were superimposed on the three-dimensional anatomical images to generate the functional and anatomical images, which were then normalized into the Montreal Neurological Institute (MNI) stereotactic space for

anatomical localization of the activated brain regions. The xjView toolbox (Version 9.7; <https://www.alivelearn.net/xjview>) was adopted to visualize cerebral activations induced by four configurations of the auditory language stimuli in the current study.

## 3. Results And Discussion

### 3.1 ERP data

Table 2 illustrates the amplitudes and latencies of N400 and P600 components elicited by the FL-R, L-FR, FL-FR, and NL-NR stimuli in three separate runs of recording. As N400 and P600 components were typically distributed in the centroparietal areas and reached the maximum at the midline electrode sites (Brouwer et al., 2012; Brouwer & Hoeks, 2013; Swaab et al., 2012; van Herten et al., 2005), analyses of the amplitudes and latencies of N400 and P600 are carried out at the midline sites of Fz, Cz, and Pz. The environmental sounds (running water, train noises, and birds singing) as novel stimuli in three runs of ERP recording did not elicit any language-related N400 and P600 effects, thus confirming the connection between N400 and P600 with language.

**Table 2.** N400 and P600 amplitudes and latencies in FL-R, L-FR, and FL-FR configurations

		N400			P600	
Configuration	Electrode	Stimulus	Amplitude ( $\mu\text{V}$ )	Latency (ms)	Amplitude ( $\mu\text{V}$ )	Latency (ms)
FL-R	Fz	NL-NR	-8.36	568	9.06	655
		FL-R	-12.61	584	6.26	705
	Cz	NL-NR	-6.23	568	8.98	638
		FL-R	-11.41	580	4.99	720
	Pz	NL-NR	-3.20	569	7.27	635
		FL-R	-10.71	564	3.64	721
L-FR	Fz	NL-NR	-2.53	450	3.10	531
		L-FR	-6.10	418	4.86	613
	Cz	NL-NR	-2.52	434	5.07	482
		L-FR	-6.26	424	5.50	633
	Pz	NL-NR	-3.14	419	3.64	500
		L-FR	-5.35	397	8.90	823
FL-FR	Fz	NL-NR	-6.67	444	6.66	546
		FL-FR	-6.30	461	6.09	728
	Cz	NL-NR	-7.22	444	7.46	534
		FL-FR	-5.28	467	2.73	731
	Pz	NL-NR	-5.98	442	2.95	532
		FL-FR	-4.07	433	3.03	750

In the first run of the experiment regarding the FL-R configuration, the amplitudes of the N400 component induced by the unfiltered NL-NR stimuli were smaller than the FL-R signals elicited at the midline sites. But for the P600 component, the FL-R stimuli elicited smaller amplitudes than the NL-NR stimuli did. The latencies of the N400 and P600 components showed that FL-R elicited longer response time compared to the latencies induced by the NL-NR signals, except for the N400 latency at the Pz site elicited by FL-R was slightly shorter. As for the L-FR configuration, the L-FR stimuli induced larger amplitudes of both N400 and P600 components in the centroparietal areas compared with NL-NR. In addition, the L-FR signals elicited shorter latencies of the N400 component but longer P600 latencies at the midline sites relative to the NL-NR stimuli. Compared to the NL-NR signals, FL-FR elicited smaller N400 amplitudes, and smaller P600 amplitudes were also induced by the FL-FR signals except for the Pz site that showed a slightly

larger amplitude. Longer latencies of N400 induced by FL-FR occurred at the Fz and Cz sites but shorter latency at the Pz site relative to NL-NR. P600 latencies reflected that FL-FR elicited a longer response time than NL-NR did.

The absolute voltage values of ERPs elicited by the FL-R, L-FR, FL-FR, and NL-NR stimuli at the midline sites (Fz, Cz, and Pz sites) are plotted in Figure 5. ERP waves are plotted in the FL-R, L-FR, and FL-FR configurations respectively corresponding to three separate runs of ERP recording. Green waves refer to the deviant stimuli (i.e., the FL-R, L-FR, and FL-FR stimuli in three runs of the experiment respectively) and blue waves represent the standard NL-NR stimuli in three runs. Since the difference waves between deviant and standard stimuli provide more information on the impact of linguistic manipulations on brain responses rather than the absolute voltage values (Morgan-Short & Tanner, 2014), the difference waves between FL-R/L-FR/FL-FR and NL-NR are displayed by red waves to identify the N400 and P600 effects.

N400 and P600 effects were elicited due to linguistic (semantic and syntactic) anomalies relative to well-formed linguistic signals (Kutas & Federmeier, 2011; Swaab et al., 2012). From the waveforms in the FL-R configuration, FL-R elicited larger N400 amplitudes compared to NL-NR, which indicates a higher processing load of semantic manipulation elicited by the FL-R signals. The N400 effect was elicited by FL-R at around 500 ms after the stimulus onset in the Pz site. But for P600, FL-R did not elicit any obvious P600 effect since smaller amplitudes were elicited at the midline sites relative to NL-NR. It indicated that a lower load of syntactic manipulation was elicited by FL-R than NL-NR. The FL-R configuration shows that the brain actively manipulates semantic processing which occupies a higher processing load compared to NL-NR, meanwhile, the FL-R signals are syntactically easier to process.

Regarding the L-FR configuration, both N400 and P600 effects were elicited by the L-FR signals at the midline sites. Different from the other configurations of FL-R and FL-FR, difference waves between L-FR and FL-FR fluctuated considerably at three midline sites. It indicates that the L-FR signals are unusual and unexpected for the brain to process, as a result, the brain struggles with the L-FR stimuli.

The waveforms in the FL-FR configuration showed that smaller N400 amplitudes were elicited by FL-FR compared to NL-NR, and no N400 effect was found in the midline sites. But the P600 effect induced by the FL-FR signals was observed at the Pz site. It indicated that the FL-FR stimuli were semantically easier for the brain to process than NL-NR since the FL-FR signals were prosodic signals without identifiable lexical information, however, the FL-FR stimuli with only intonation and rhythm information were unusual for the participant to manipulate syntactic processing.

Topographic maps of the ERP components N400 and P600 that were elicited by the standard and deviant stimuli in three runs of the experiment were plotted in Figure 6. Since the novel stimuli of environmental sounds did not elicit the N400 and P600 components during three runs of recording, the scalp topographies of ERP components elicited by the environmental sounds were not plotted in Figure 6. In the

topographic maps, the N400 component with negative voltage values was presented on the left, and the positive P600 component was shown on the right.

As standard stimuli in the experiment, the unfiltered NL-NR stimuli maintained all frequencies of the auditory signals so that each word in the sentences could be identified. Thus, a smaller processing load and shorter response time for semantic manipulation were elicited by the NL-NR signals relative to the FL-R stimuli. But it showed a quite small mental workload for syntactic processing of the FL-R signals though response time to FL-R was longer compared to NL-NR. Topographic maps showed that the N400 and P600 components elicited by NL-NR were distributed in central and frontal areas symmetrically, but FL-R induced left-lateralized patterns for semantic and syntactic processing in the frontal area peaking at 584 ms and 719 ms respectively.

The L-FR signals elicited larger amplitudes of N400 and P600 components, which indicated that a heavier mental load was required for semantic and syntactic manipulations relative to the NL-NR signals. This may result from the L-FR signals violating the left ear advantage for prosodic information and the right ear advantage for linguistic signals (Meyer et al., 2002; Sammler et al., 2015; Tervaniemi & Hugdahl, 2003; Vigneau et al., 2006). L-FR sent prosodic signals to the right ear and linguistic information to the left ear, which is presumably a non-optimal signal for the brain to process. As to latencies, a shorter response time elicited by L-FR was observed for semantic processing, but a longer response time was detected for syntactic manipulation compared to the NL-NR signals. Topographic maps indicated that NL-NR in the second run of the experiment elicited N400 and P600 with general symmetrical distributions in the central areas, similar to the distributions in the first run of the experiment. In addition, the N400 component elicited by the L-FR signals was distributed in the central and frontal areas, and the P600 component was detected in the centroparietal area. It is not obvious that L-FR elicited lateralization during semantic and syntactic processing.

For both-ears-filtered FL-FR stimuli, smaller amplitudes of N400 were elicited by FL-FR, indicating a small processing load for semantic manipulation. Further, smaller P600 amplitudes elicited by FL-FR were also found in the central and frontal areas, which suggested a lower mental workload for syntactic processing. This result supports the assumptions of the authors that 320 Hz low-pass filtering reduces the processing load for language signals since the filtered sounds only contain intonation and rhythm information that lightens the load for processing meanings of the signals (Asp et al., 2012; Guberina, 1972; He et al., 2015; Lian, 1980; Yang, 2016). As to response time, FL-FR with only prosodic information sounded unusual and was an unexpected language signal for the participant, and thus, required a longer response time to process both for semantic and syntactic information. Only one exception occurred in the centroparietal area where FL-FR induced a shorter response time regarding semantic processing relative to the NL-NR signals. Topographic maps suggest that N400 and P600 elicited by NL-NR presented a basically symmetrical distribution in the central and frontal areas, which was consistent with the results of distributions in the other two runs of the experiment. But the N400 component elicited by FL-FR was

distributed in the occipital area and the P600 showed a slightly left-lateralized distribution in the central area.

Additionally, the environmental sounds of running water, train noises, and birds singing, used as novel stimuli in three runs of the experiment, did not elicit any language-related N400 or P600 component. It indicates that the brain distinguishes language and non-language signals successfully and processes them differently.

### **3.2 fMRI data**

To investigate brain regions for processing low-pass filtered and unfiltered stimuli under dichotic listening conditions, four runs of fMRI scanning with a rest-stimulus block design were performed for the FL-FR, FL-R, L-FR, and NL-NR signals respectively. Brain activations and activation maps for FL-FR, FL-R, L-FR, and NL-NR stimuli were presented in Table 3 and Figure 7.

**Table 3.** Brain activations for FL-FR, FL-R, L-FR, and NL-NR stimuli

Stimulus	Region	BA	k	z-value	T	MNI coordinates		
						x	y	z
FL-FR <sup>a</sup>	L-MFG	-	158	3.69	3.92	-40	48	-6
	R-IFG	-	113	3.38	3.56	50	42	4
		46		3.33	3.50	54	44	18
		9		3.61	3.82	62	16	32
	L-Putamen	-	111	3.14	3.29	-14	16	6
	L-STG	-	110	3.61	3.83	-68	-22	12
		42		3.60	3.82	-70	-32	20
	R-SFG	-	101	3.69	3.92	-40	48	-6
		10		3.49	3.69	24	62	-6
	R-STG	22	97	3.71	3.94	54	-12	8
	R-Cun	-	93	3.42	3.60	12	-88	36
	R-MTG	-	55	3.27	3.43	40	-70	26
	FL-R <sup>a</sup>	L-Inferior parietal lobule	-	871	4.2	4.54	-34	-60
40				4.17	4.51	-44	-52	-60
42				3.90	4.17	-70	-22	14
L-MFG		8	195	3.73	3.97	-26	24	62
		10	80	3.60	3.82	-39	45	30
R-Inferior parietal lobule		40	180	3.37	3.55	60	-50	42
R-Precuneus		7	74	3.45	3.65	2	-62	52
R-STG		22	56	3.65	3.88	60	-6	6
R-MFG		10	51	3.62	3.84	36	42	24
L-STG		41	19	3.26	3.43	-42	-34	12
L-FR <sup>b</sup>		Corpus callosum	-	475	2.14	2.19	6	-18
	R-STG	-	356	3.05	3.19	68	-6	-8
		22		2.72	2.82	58	-8	8

	Midbrain	-	103	2.10	2.15	4	-16	-16
	L-Brainstem	-	65	2.41	2.48	-16	-22	-14
	R-MFG	6	55	2.71	2.80	60	4	52
	L-STG	41	53	2.58	2.66	-40	-24	10
		42	41	2.32	2.38	-58	-16	10
NL-NR <sup>a</sup>	L-STG	-	1364	4.83	5.35	-70	-30	18
		42		4.82	5.34	-60	20	12
		40		4.82	5.34	-60	-20	12
	R-STG	-	796	4.92	5.48	58	6	54
		22		4.22	4.57	66	-6	8
		43		4.53	4.96	54	-10	10
	L-Precentral gyrus	6	81	3.69	3.92	-52	-4	56
	L-Pons	-	69	3.48	3.68	-10	-34	-38
	L-Cerebellum anterior lobe	-	56	3.33	3.50	-8	-50	-24
	R-PoCG	2	47	4.17	4.50	62	-30	54
	R-Pons	-	18	3.29	3.45	14	-32	-36
	Midbrain	-	15	3.30	3.47	8	-18	-18

Notes: BA, Brodmann area;  $k$ , cluster size (number of voxels); L, left hemisphere; R, right hemisphere; MFG, middle frontal gyrus; IFG, inferior frontal gyrus; STG, superior temporal gyrus; SFG, superior frontal gyrus; Cun, cuneus; MTG, middle temporal gyrus; PoCG, postcentral gyrus.

<sup>a</sup> Clusters are thresholded at  $p < .001$  (uncorrected).

<sup>b</sup> Clusters are thresholded at  $p < .03$  (uncorrected).

The FL-FR signals induced involvement of both hemispheres, including the left middle frontal gyrus (MFG), superior temporal gyrus (STG), and putamen. In the right hemisphere, the inferior frontal gyrus (IFG), superior frontal gyrus (SFG); STG, Cuneus (Cun); and middle temporal gyrus (MTG) were activated as well.

For FL-R, stronger activity was found in the left hemisphere, in which the inferior parietal lobule (BA 40 and 42), MFG, and STG (Heschl's gyrus, BA 41) were activated. Increased activation was also observed in

the right hemisphere, including the areas of the inferior parietal lobule (BA 40), precuneus, STG, and MFG.

L-FR led to relatively lower activation levels compared to the other three stimuli, which did not reveal significant activation at the cluster threshold of  $p < .001$ . When the threshold of the  $p$ -cluster was set at  $< .03$ , increased activation was detected in the right STG and MFG, and the left brainstem and STG. However, the corpus callosum was more activated than the above-mentioned areas.

As to the NL-NR signals, activation of bilateral STG was the strongest relative to the activation levels induced by the other three stimuli. In the left hemisphere, increased activation was found in STG (BA 42 and 40), precentral gyrus, pons, and cerebellum anterior lobe. The right STG (BA 22 and 43), postcentral gyrus (PoCG), and pons were detected with increased activation. Activation of the midbrain was detected as well.

fMRI results indicate that the neural processing patterns of low-pass filtered and unfiltered language signals do differ in the four dichotic configurations of stimuli (i.e., FL-FR, FL-R, L-FR, and NL-NR stimuli). Even though the English sentences as listening materials were the same in four runs of the experiment, the only difference of dichotic listening to filtered and unfiltered sentences induced different activation patterns of processing. Generally, 320 Hz low-pass filtered signals led to lower activation levels in the brain. The both-ears-filtered stimuli FL-FR and the unfiltered sentence signals (NL-NR) induced increased activation in the regions of bilateral STG, but NL-NR induced more activated regions extended to bilateral inferior parietal lobule and the right precentral gyrus, further, the midbrain, cerebellum, pons, posterior cingulate, and corpus callosum were increasingly activated by NL-NR. In general, higher levels of activation were induced by the NL-NR signals relative to FL-FR. It confirms the previous assumption by the authors that low-pass filtered language signals could lighten the listener's mental workload for processing the meanings of words (Asp et al., 2012; Guberina, 1972; He et al., 2015; Lian, 1980; Yang, 2016).

FL-FR, as a signal of speech intonation and rhythm, presents a neural processing pattern of linguistic prosody. It revealed that left MFG, bilateral STG, right IFG, and SFG were involved in prosodic processing, in addition, more frontal areas of the right hemisphere got involved. This result is consistent with the findings of the recent studies conducted by Chien et al. (2020, 2021) that Chinese (a tonal language) speakers recruit bilateral fronto-temporal regions for intonation processing. By using the unfiltered monosyllabic Mandarin words as auditory stimuli, Chien et al. (2020, 2021) proposed that the connection between left IFG and bilateral temporal areas may reflect the phonological processing network for auditory intonation perception and prosodic categorization. However, the current study revealed stronger activation in right IFG induced by FL-FR. This may result from the stimuli used in the current study. Unlike the phonological processing of the unfiltered words, the low-pass filtered sentence signals contain only prosodic information without noticeable segmental features of the speech sounds. Thus, the FL-FR signals induced activation in the bilateral fronto-temporal areas for intonation and rhythm processing, in addition, a pathway linking posterior temporal to IFG in the right hemisphere was prominently involved in the prosodic processing (Meyer et al., 2002; Sammler et al., 2015).

The FL-R stimuli that are assumed to be consistent with ear advantage sent prosodic signals to the right hemisphere and directed linguistic signals to the left hemisphere. It revealed a left-dominant processing pattern that the left hemisphere was actively involved in the processing of the unfiltered signals, in the meantime, the activated areas were smaller and the activation levels were lowered by the filtered signals in the right hemisphere. Similar to the results in FL-FR, this result may suggest that the low-pass filtered stimuli, as discussed above, do induce lower activation levels and smaller activated areas in the right hemisphere. Whether the filtered prosodic signals were sent to the brain through both-ears-filtered configuration (FL-FR) or dichotically (FL-R), low-pass filtering induced activation of the right-hemispheric STG and MFG, indicating an auditory prosodic processing pattern with a lower activation level relative to the unfiltered signal. As to the linguistic signal dichotically sent to the left hemisphere, stronger involvement of left inferior parietal lobule (supramarginal gyrus, BA 40, and posterior transverse temporal area, BA 42), Heschl's gyrus (BA 41), and MFG were detected. This result is basically consistent with the "semantic hubs" for manipulating a concept or meaning of the spoken or written language symbols (Pulvermüller, 2013; Pulvermüller & Fadiga, 2016), including left-hemispheric inferior frontal areas (Bookheimer, 2002), inferior parietal regions (Binder & Desai, 2011), anterior or posterior-middle temporal areas (Hickok & Poeppel, 2007; Patterson et al., 2007). Although left inferior frontal regions in the "semantic hubs" were not significantly activated by the FL-R signals, increased activation was observed in the middle frontal gyrus. This result may be led by the L2 auditory stimuli that MFG is related to higher-level cognitive control and essential for effective communication together with inferior frontal regions especially in a foreign language (Mårtensson et al., 2012; Sierpowska et al., 2018). Thus, FL-R induces prosodic and semantic processing patterns with lower activation levels and small activated areas in the right hemisphere. Further, the left-hemispheric semantic processing pattern and higher-level cognitive processing of foreign language signals are significantly induced by the FL-R signals.

In terms of L-FR, significant activation was shown at the  $p$ -cluster threshold  $< .03$  rather than  $p$ -cluster  $< .001$  in the other three stimuli. It indicated that a relatively lower activation level was induced by L-FR in the whole brain. In other words, L-FR could not induce as significant activation or active involvement as the other three signals. According to ear advantage, L-FR is presumably a non-optimal language input that violates the left and right ear advantages, which may be the reason that both hemispheres are reluctant to process the signal. Activation of the primary auditory cortex in bilateral STG especially Heschl's gyrus (Da Costa et al., 2011) was revealed, but not a specific language processing pattern. In addition, greater activation was found in the corpus callosum induced by the L-FR signals, even greater than other cerebral regions. The corpus callosum, connecting both hemispheres, was identified as assisting language processing and language lateralization (Hinkley et al., 2016). As L-FR is not favored by both hemispheres, stronger activation of the corpus callosum may indicate that the left and right hemispheres redirect the signal to the other hemisphere after they receive the signal. As a result, L-FR leads to activation of the primary auditory cortex and lower activation levels in the brain. However, both hemispheres seem to send the signals away instead of processing them.

### **3.3 Relationship between ERP and fMRI results**

ERP and fMRI results revealed distinct neural processing patterns of the FL-FR, FL-R, L-FR, and NL-NR stimuli, which supported our assumption that dichotic listening to the low-pass filtered prosodic and unfiltered linguistic signals induces different processing mechanisms due to the left and right ear advantages. As the English (L2) sentences as auditory stimuli in each run of the combined ERP and fMRI experiment were the same, the differences in processing patterns were led by the dichotic configurations of the filtered and unfiltered stimuli. Thus, there should be an optimal or non-optimal language signal for foreign language learners to process, which are consistent with or violate the left and right ear advantages.

Compared with NL-NR, FL-FR in ERP and fMRI experiments showed lower mental load for processing. Amplitudes of the ERP components regarding semantic and syntactic processing elicited by FL-FR suggested that FL-FR released the processing load compared with the unfiltered NL-NR stimuli, although response time for semantic and syntactic manipulations was longer. The reason for the longer response time may be that the filtered stimuli are limited in semantic and syntactic information and unexpected for the participant to process. fMRI results revealed lower levels of activation compared to NL-NR, in addition, a neural processing pattern for prosody was detected in the bilateral fronto-temporal areas, especially in the right hemisphere. Topographic maps obtained from the ERP experiment showed that FL-FR induced a slightly left-lateralized distribution in the central areas regarding syntactic processing, which was basically consistent with the fMRI results that stronger activation was found in the left middle frontal gyrus.

The brain was actively involved in semantic manipulation of the FL-R signals with higher mental load and longer response time in the ERP results relative to NL-NR. But it was syntactically easier to process FL-R with a lower mental workload. A left-lateralized distribution of the components was identified in the ERP experiment, which was confirmed in the fMRI experiment. fMRI results showed stronger involvement of the left hemisphere with prosodic and semantic processing patterns. The “semantic hubs” in the left hemisphere (Pulvermüller, 2013; Pulvermüller & Fadiga, 2016) were activated by the FL-R signals with a higher level of cognitive control of L2 signals. Meanwhile, the right hemisphere showed lower activation levels with smaller areas. FL-R, consistent with the left and right ear advantages, induced lower mental workload in the right hemisphere by sending prosodic information, at the same time, linguistic signals sent to the left hemisphere led to stronger activation for semantic processing. The higher mental load and longer response time detected in the ERP experiment may result from the foreign language signals as stimuli, which required a higher level of cognitive control due to significant activation of left MFG as observed in the fMRI experiment (Mårtensson et al., 2012; Sierpowska et al., 2018). FL-R can be identified as an optimal language input because prosodic and semantic processing patterns can be significantly activated in the left hemisphere, and the activation level or processing load in the right hemisphere was lowered. FL-R optimized the auditory language input by actively involving left-hemispheric dominance for language and lowering cognitive load in the right hemisphere for other higher-order/complex cognitive processes (Galotti, 2017; Levine, 2009).

L-FR induced a higher processing load for semantic and syntactic manipulations in the ERP experiment. Shorter response time for semantic processing and longer response time for syntactic manipulation were observed compared to the NL-NR signals. The N400 and P600 effects elicited by L-FR indicated that the brain struggled to process this signal. Different from the higher processing load detected in the ERP experiment, fMRI results suggested a quite smaller activation level relative to the other three stimuli. In addition, both hemispheres were struggling with the signals that they received and were reluctant to process, apparently redirecting the signals to the other hemisphere via the corpus callosum. As a result, stronger activation induced by L-FR was found in the corpus callosum rather than other cerebral areas. Thus, L-FR can be identified as a non-optimal auditory language input for language learners.

The ERP study also revealed that the neural mechanisms for processing language signals and environmental sounds (non-verbal signals) were different, so environmental sounds did not trigger any language-related ERP components.

## 4. Conclusion

To sum up, this case study employed a combined ERP and fMRI experiment to identify an optimal auditory language signal for Chinese ESL learners. The FL-R signal derived from principles of verbotonalism and used in a dichotic context can be considered as the optimal signal for Chinese ESL learners as it is demonstrably consistent with the left and right ear advantages. Specifically, FL-R induced active prosodic and semantic processing of the linguistic signals in the left hemisphere and, in the meantime, lowered the processing load in the right hemisphere. The optimized FL-R language input signal appears best suited for optimal processing. As a result, FL-R should, in principle, enable learners to make better sense of foreign language signals and to facilitate language learning (Guberina & Asp, 1981, 2013; Lian, 2004). In addition, the current study also identified a non-optimal language signal L-FR, which should be avoided in learning or teaching a language. While the authors acknowledge that this is a limited case study requiring additional confirmation, the current study finds its origins in a rater-based study. Results showed that the group of participants listening to the FL-R signals during an ESL course outperformed the L-FR and NL-NR groups in English pronunciation, especially in intonation and fluency performance. At the same time, the L-FR group achieved the least improvement. As a result, it is concluded that there does exist an optimal language input (FL-R) and a non-optimal input (L-FR) for foreign language learners. A clear pedagogical implication is that ESL learners (perhaps others too) would benefit from dichotic FL-R input while learning English. Future studies should recruit a larger number of participants to generalize the findings.

## Abbreviations

BA	Brodman area
BOLD	Blood oxygenation level dependent
Cun	Cuneus
ERP	Event-related potential
ESL	English as a second language
F <sub>0</sub>	Fundamental frequency
FL-FR	Filtered in both left and right channels
FL-R	Filtered in the left channel and unfiltered in the right channel
fMRI	Functional magnetic resonance imaging
Hz	Hertz
IFG	Inferior frontal gyrus
<i>k</i>	Cluster size (number of voxels)
L/LH	Left hemisphere
L1	The first language (Mandarin Chinese in this study)
L2	The second language (English in this study)
L-FR	Unfiltered in the left channel and filtered in the right channel
MFG	Middle frontal gyrus
ms	Millisecond
MTG	Middle temporal gyrus
NL-NR	Unfiltered in both left and right channels
PoCG	Postcentral gyrus
R/RH	Right hemisphere
s	Second
SFG	Superior frontal gyrus
STG	Superior temporal gyrus

## Declarations

- Availability of data and material

Please contact the corresponding authors for data requests.

- Funding

This work was supported by Suranaree University of Technology, Thailand under Grant No. SUT2-203-57-12-01; Scientific Research Projects of Yunnan Provincial Department of Education, China under Grant No. 2018JS180; and Scientific Research Projects of Yunnan Provincial Science and Technology Department, China under Grant No. 2019FE001(-214).

- Competing interests

The authors declare that they have no competing financial interests.

- Authors' contributions

AL, XC, and NP conceived and designed the study. HC, JO, and WZ carried out the ERP experiment. WZ performed ERP statistical analysis. YM and YZ designed the fMRI experiment. YS carried out the fMRI experiment and analyzed fMRI data. XC and AL wrote the paper. AL reviewed and edited the manuscript. All authors read and approved the final manuscript.

- Acknowledgements

Not applicable

## References

1. Antonenko, P., van Gog, T., & Paas, F. (2014). Implications of Neuroimaging for Educational Research. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 51–63). Springer. [https://doi.org/10.1007/978-1-4614-3185-5\\_5](https://doi.org/10.1007/978-1-4614-3185-5_5).
2. Asp, C. W. (2006). *Verbotonal speech treatment*. Plural Publishing.
3. Asp, C. W., Kim, Y., & Davis, C. (2003). *Verbotonal approach for the children with cochlear implant in public schools. I. Congresso dell'Associazione Internazionale per il Metodo Verbotonale*. Rome: Neuroscienze e Metodo Verbotonale.
4. Asp, C. W., Kline, M., & Koike, K. J. (2012). Verbotonal worldwide. In R. Goldfarb (Ed.), *Translational speech-language pathology and audiology: Essays in honor of Dr. Sadanand Singh* (pp. 319–326). Plural Publishing.
5. Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>.
6. Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, 25, 151–188.
7. Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the role of the P600 in language comprehension. *Brain research*, 1446, 127–143.

8. Brouwer, H., & Hoeks, J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*, 1–12.
9. Chien, P. J., Friederici, A. D., Hartwigsen, G., & Sammler, D. (2020). Neural correlates of intonation and lexical tone in tonal and non-tonal language speakers. *Human brain mapping*, *41*(7), 1842–1858. <https://doi.org/10.1002/hbm.24916>.
10. Chien, P. J., Friederici, A. D., Hartwigsen, G., & Sammler, D. (2021). Intonation processing increases task-specific fronto-temporal connectivity in tonal language speakers. *Human brain mapping*, *42*, 161–174. <https://doi.org/10.1002/hbm.25214>.
11. Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S., Clarke, S., & Saenz, M. (2011). Human primary auditory cortex follows the shape of Heschl's gyrus. *Journal of Neuroscience*, *31*(40), 14067–14075. <https://doi.org/10.1523/JNEUROSCI.2000-11.2011>.
12. Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in human neuroscience*, *8*, 437.
13. Donaldson, D. I. (2004). Parsing brain activity with fMRI and mixed designs: What kind of a state is neuroimaging in? *Trends in Neurosciences*, *27*, 442–444. <https://doi.org/10.1016/j.tins.2004.06.001>.
14. Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*(3), 165–189.
15. Galotti, K. M. (2017). *Cognitive development: Infancy through adolescence* (2nd ed.). Sage.
16. Guberina, P., & Asp, C. W. (1981). *The verbo-tonal method for rehabilitating people with communication problems*. World Rehabilitation Fund.
17. Guberina, P., & Asp, C. W. (2013). *The Verbotonal Method*. Artresor Naklada.
18. Guberina, P. (1972). *Restricted bands of frequencies in auditory rehabilitation of deaf*. Institute of Phonetics, Faculty of Arts, University of Zagreb.
19. Hashemi, L., & Thomas, B. (2013). *Objective PET Teacher's Book* (4th, ed.). Cambridge University Press.
20. He, B. (2014). *Improving the English Pronunciation of Chinese EFL Learners through the Integration of CALL and Verbotonalism* [Doctoral dissertation, Suranaree University of Technology]. Thailand.
21. He, B., Sangarun, P., & Lian, A. P. (2015). Improving the English pronunciation of Chinese EFL university students through the integration of CALL and verbotonalism. Seventeenth international CALL research conference: Task design and CALL, Tarragona, Spain: University of Antwerp.
22. Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. <https://doi.org/10.1038/nrn2113>.
23. Hinkley, L. B. N., Marco, E. J., Brown, E. G., Bukshpun, P., Gold, J., Hill, S., Findlay, A. M., Jeremy, R. J., Wakahiro, M. L., Barkovich, A. J., Mukherjee, P., Sherr, E. H., & Nagarajan, S. S. (2016). The contribution of the corpus callosum to language lateralization. *Journal of Neuroscience*, *36*(16), 4522–4533. <https://doi.org/10.1523/JNEUROSCI.3850-14.2016>.

24. Indefrey, P. (2012). Hemodynamic Studies of Syntactic Processing. In M. Faust (Ed.), *The Handbook of Neuropsychology of Language* (pp. 209–228). Wiley-Blackwell.
25. Jurjević-Grkinić, I., Munivrana, B., & Mijić, V. (2015). Applying a Verbotonal Method to Rehabilitation of the Hearing Impaired Children. Symposium & 52nd Inner Ear Biology Workshop, Pisa, Italy.
26. Kim, Y., & Asp, C. W. (2002). Low Frequency Perception of Rhythm and Intonation Speech Patterns by Normal Hearing Adults. *Speech Sciences*, 9(1), 7–16.
27. Klem, G. H., Lüders, H. O., Jasper, H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology. Supplement*, 52(3), 3–6.
28. Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
29. Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
30. Krashen, S. D. (2003). *Explorations in language acquisition and use*. Heinemann.
31. Kuhl, P. K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience*, 31, 511–534.
32. Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
33. Levine, M. D. (2009). Differences in Learning and Neurodevelopmental Function in School-age Children. In W. B. Carey, A. C. Crocker, W. L. Coleman, E. R. Elias & H. M. Feldman (Eds.), *Developmental-Behavioral Pediatrics* (4th ed., pp. 535–546). W.B. Saunders.  
<https://doi.org/10.1016/B978-1-4160-3370-7.00055-9>.
34. Lian, A. P. (1980). *Intonation patterns of French (Teacher's book)*. River Seine Publications Pty Ltd.
35. Lian, A. P. (2004). Technology-enhanced language learning environments: A rhizomatic approach. In J.-B. Son (Ed.), *Computer-assisted language learning: Concepts, contexts and practices* (pp. 1–20). iUniverse.
36. Lian, A. P. (2011). Reflections on language-learning in the 21st century: The rhizome at work. *Rangsit Journal of Arts and Sciences*, 1(1), 5–17.
37. Lian, A. P. (2014). On-demand generation of individualized language learning lessons. *Journal of Science*, 9(1), 25–38.
38. Lian, A. P., & Sussex, R. (2018). Toward a Critical Epistemology for Learning Languages and Cultures in Twenty-First Century Asia. In A. Curtis & R. Sussex (Eds.), *Intercultural Communication in Asia: Education, Language and Values* (pp. 37–54). Springer International Publishing.  
[https://doi.org/10.1007/978-3-319-69995-0\\_3](https://doi.org/10.1007/978-3-319-69995-0_3).
39. Logothetis, N. K. (2012). What We Can and What We Can't Do with fMRI. Society for Neuroscience: 2012 Short Course II MRI and Advanced Imaging in Animals and Humans, New Orleans, LA.
40. Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.
41. Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *Neuroimage*, 63(1),

240–244.

42. McLaughlin, B. (1987). *Theories of second-language learning*. Edward Arnold.
43. Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human brain mapping, 17*(2), 73–88. <https://doi.org/10.1002/hbm.10042>.
44. Morgan-Short, K., & Tanner, D. (2014). Event-related potentials (ERPs). In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 127–152). Routledge.
45. Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.
46. Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976–987. <https://doi.org/10.1038/nrn2277>.
47. Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: The analysis of functional brain images*. Elsevier.
48. Petersen, S. E., & Dubis, J. W. (2012). The mixed block/event-related design. *Neuroimage, 62*(2), 1177–1184. <https://doi.org/10.1016/j.neuroimage.2011.09.084>.
49. Poldrack, R. A. (2018). *The new mind readers: What neuroimaging can and cannot reveal about our thoughts*. Princeton University Press.
50. Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences, 17*(9), 458–470. <https://doi.org/10.1016/j.tics.2013.06.004>.
51. Pulvermüller, F., & Fadiga, L. (2016). Brain language mechanisms built on action and perception. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 311–324). Academic Press.
52. Sammler, D., Grosbras, M. H., Anwander, A., Bestelmeyer, P. E., & Belin, P. (2015). Dorsal and ventral pathways for prosody. *Current Biology, 25*(23), 3079–3085.
53. Sierpowska, J., Fernandez-Coello, A., Gomez-Andres, A., Camins, À, Castañer, S., Juncadella, M., Gabarrós, A., & Rodríguez-Fornells, A. (2018). Involvement of the middle frontal gyrus in language switching as revealed by electrical stimulation mapping and functional magnetic resonance imaging in bilingual brain tumor patients. *Cortex; a journal devoted to the study of the nervous system and behavior, 99*, 78–92. <https://doi.org/10.1016/j.cortex.2017.10.017>.
54. Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components* (pp. 397–440). Oxford University Press.
55. Tervaniemi, M., & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews, 43*(3), 231–246.
56. van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22*(2), 241–255.

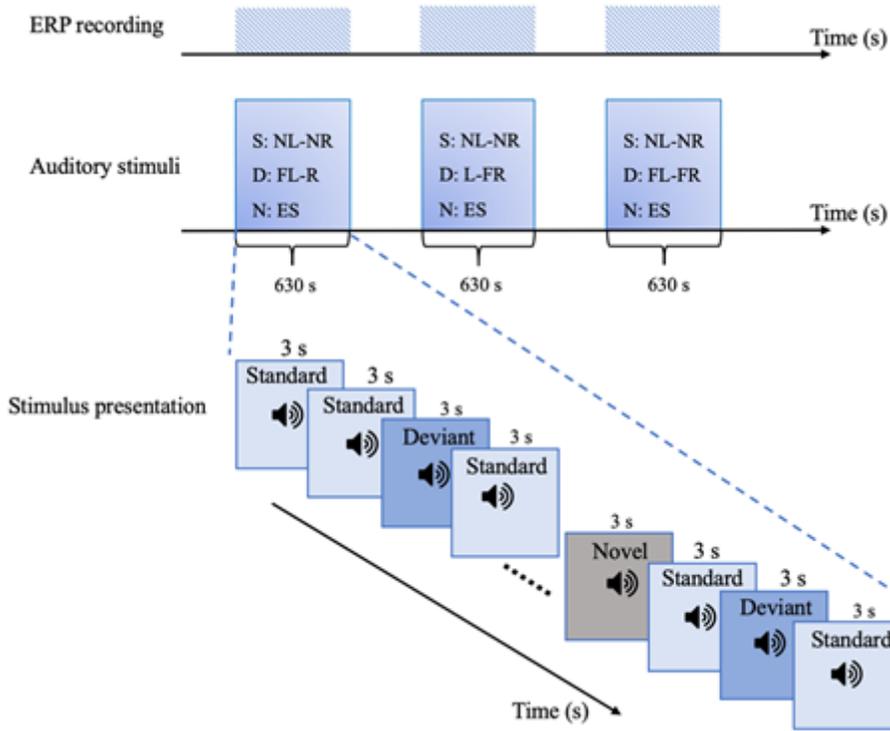
57. Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, 30(4), 1414–1432.
58. Wen, F., Lian, A. P., & Sangarun, P. (2020). Determination of corrective optimals for Chinese university learners of English. *Govor/Speech*, 37(1), 3–29.
59. Yang, Y. (2016). *Improving the Phonological Working Memory and English Speaking Skills of Chinese Primary EFL Learners with a Verbotonal-Based Approach* [Doctoral dissertation, Suranaree University of Technology]. Thailand.
60. Yang, Y., Wannaruk, A., & Lian, A. P. (2017). Improving the English-speaking skills of Chinese primary EFL learners with a verbotonal approach. *Rangsit Journal of Arts and Sciences*, 7(2), 141–156.

## Figures



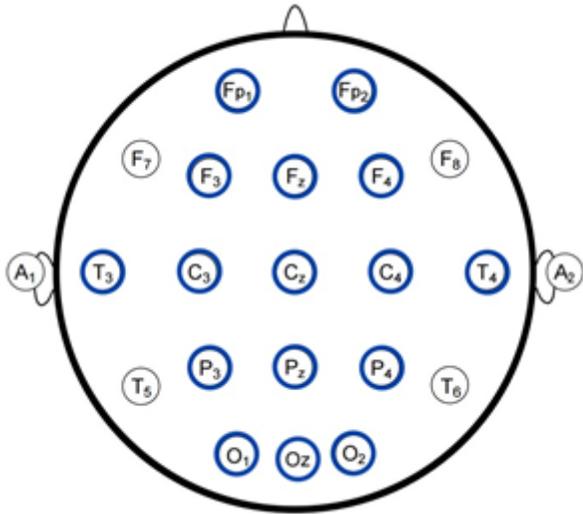
Figure 1

Spectrogram of the FL-R stimulus “I preferred baseball when I was at school.”



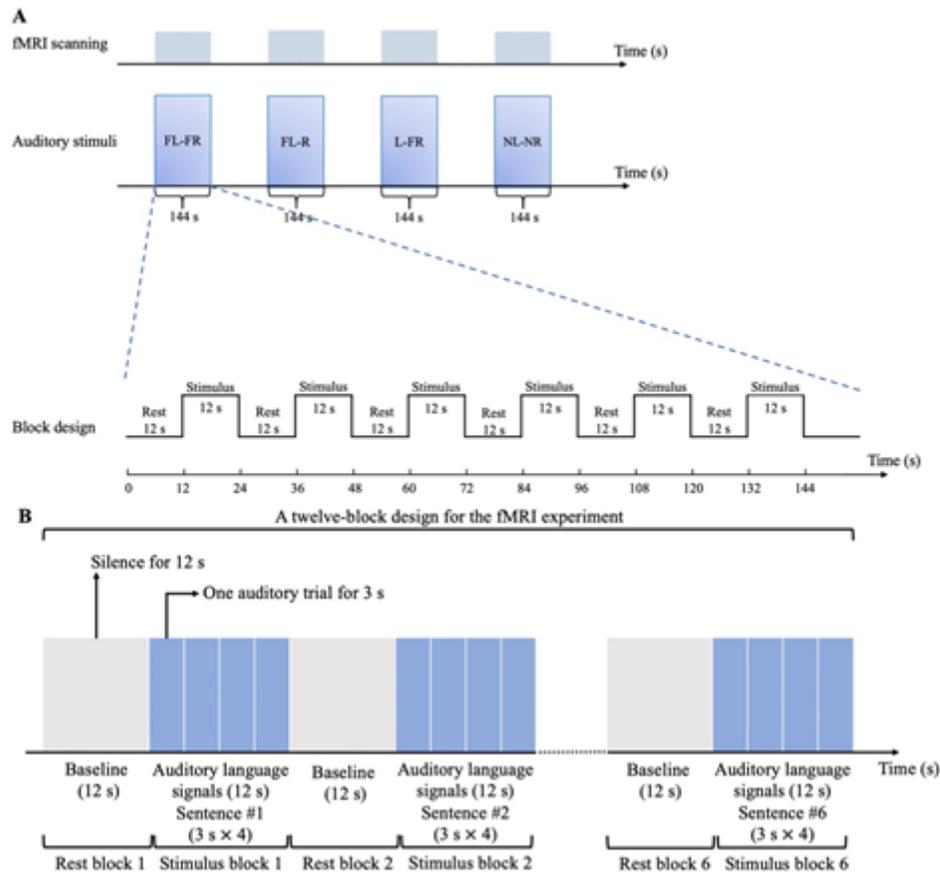
**Figure 2**

Notes: S, standard stimuli; D, deviant stimuli; N, novel stimuli; NL-NR, unfiltered stimuli in both channels; FL-R, filtered stimuli in the left channel and unfiltered stimuli in the right channel; L-FR, unfiltered stimuli in the left channel and filtered in the right channel; FL-FR, filtered stimuli in both channels; ES, environmental sounds. ERP experimental design



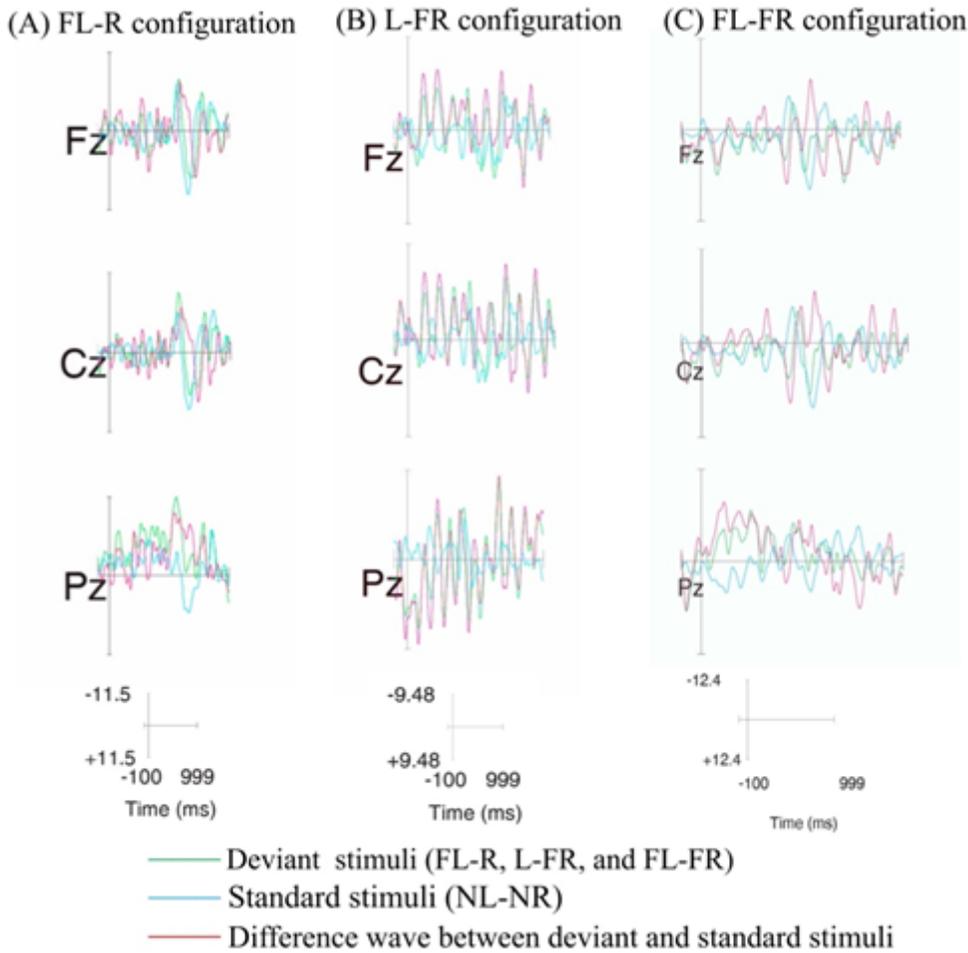
**Figure 3**

Location of electrode sites



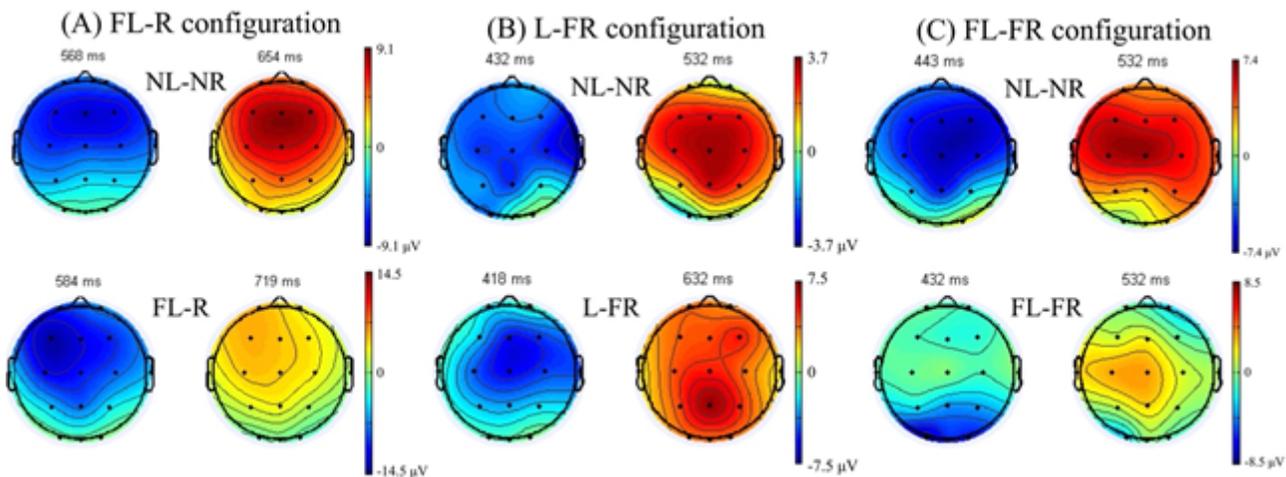
**Figure 4**

fMRI experimental design. (A) A twelve-block design for the fMRI experiment. (B) The rest-stimulus block design for the fMRI experiment.



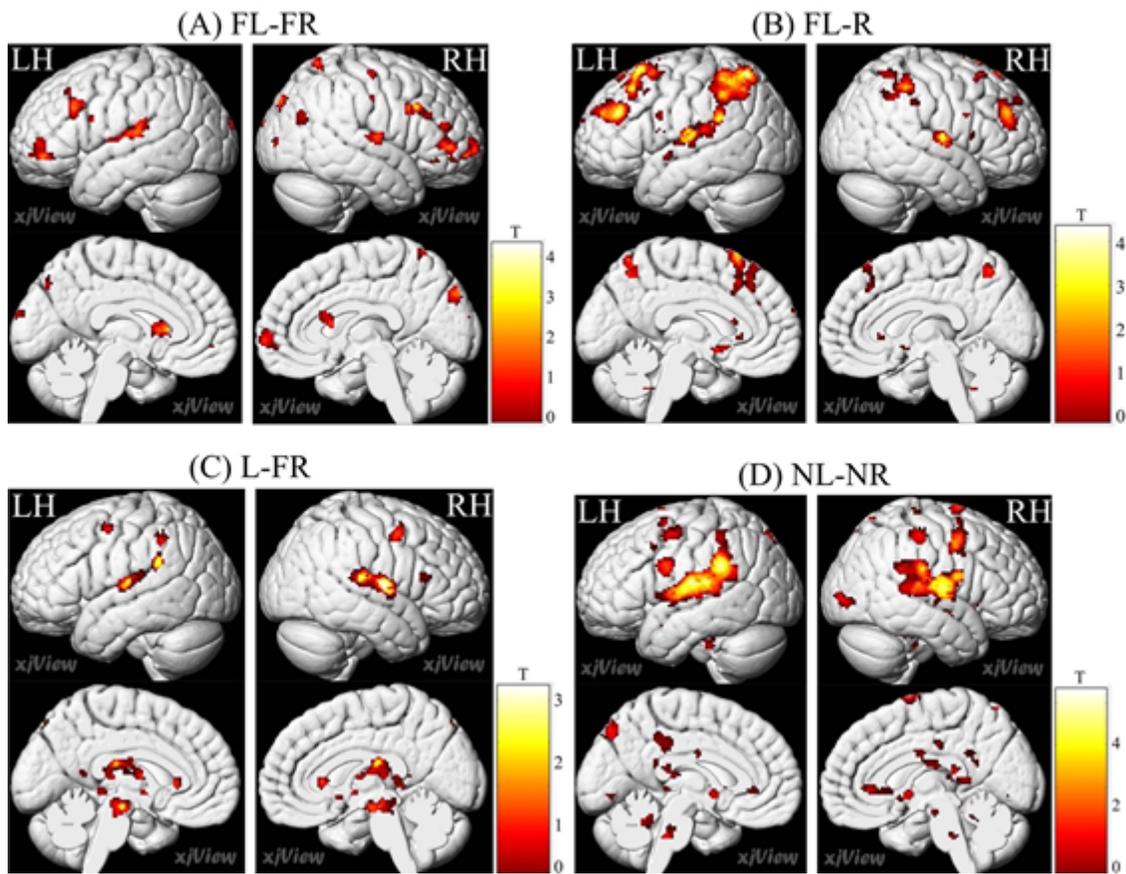
**Figure 5**

Grand ERP averages for midline sites in (A) FL-R, (B) L-FR, and (C) FL-FR configurations, and difference waves between deviant and standard stimuli



**Figure 6**

Topographic maps of the N400 and P600 components elicited by standard and deviant stimuli in (A) FL-R, (B) L-FR, and (C) FL-FR configurations



**Figure 7**

Brain activation maps for (A) FL-FR (clusters are thresholded at  $p < .001$ , uncorrected), (B) FL-R (clusters are thresholded at  $p < .001$ , uncorrected), (C) L-FR (clusters are thresholded at  $p < .03$ , uncorrected), and (D) NL-NR stimuli (clusters are thresholded at  $p < .001$ , uncorrected). LH, left hemisphere; RH, right hemisphere.