frontiers in Behavioral Neuroscience

A Proof of Concept Study of Function-based Statistical Analysis of fNIRS Data: Syntax Comprehension in Children with Specific Language Impairment Compared To Typically-Developing Controls

Guifang Fu[2], Nicholas J. Wan[3], Joseph M. Baker[1], James Montgomery[4], Julia L. Evans[5], Ronald Gillam[6*]

[1]Psychiatry, Stanford Univeristy, USA, [2]Department of Mathematics and Statistics, Utah State University, USA, [3]Department of Psychology, Utah State University, USA, [4]Communication Sciences and Disorders Programs, Ohio University, USA, [5]School of Behavioral and Brain Sciences, University of Texas at Dallas, USA, [6]Department of Communicative Disorders, Utah State University, USA

Submitted to Journal:
Frontiers in Behavioral Neuroscience

ISSN:
1662-5153

Article type:
Original Research Article

Received on:
01 Feb 2016

Accepted on:
19 May 2016

Provisional PDF published on:
19 May 2016

Frontiers website link:
www.frontiersin.org

Citation:
Fu G, Wan NJ, Baker JM, Montgomery J, Evans JL and Gillam R(2016) A Proof of Concept Study of Function-based Statistical Analysis of fNIRS Data: Syntax Comprehension in Children with Specific Language Impairment Compared To Typically-Developing Controls. *Front. Behav. Neurosci.* 10:108. doi:10.3389/fnbeh.2016.00108

frontiers

A Proof of Concept Study of Function-based Statistical Analysis of fNIRS Data: Syntax Comprehension in Children with Specific Language Impairment Compared To Typically-Developing Controls

Guifang Fu [1], **Nicholas Wan** [2], **Joseph M. Baker** [3], **James W. Montgomery** [4], **Julia L. Evans** [5], **Ronald B. Gillam** [6,*]

[1] *Department of Mathematics and Statistics, Utah State University, Logan, UT, USA 84322*
[2] *Department of Psychology, Utah State University, Logan, UT, USA 84322*
[3] *Center for Interdisciplinary Brain Sciences Research, Division of Brain Sciences, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA 94305*
[4] *Department of Communication Disorders, Ohio University, Athens, OH, USA 45701*
[5] *School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX, USA 75080*
[6] *Department of Communicative Disorders and Deaf Education, Utah State University, Logan UT, USA 84322*

Correspondence*:
Dr. Ronald B. Gillam
Department of Communicative Disorders and Deaf Education, Utah State University, Logan UT, USA 84322, ron.gillam@usu.edu

ABSTRACT

Functional near infrared spectroscopy (fNIRS) is a neuroimaging techonology that enables investigators to indirectly monitor brain activity in vivo through relative changes in the concentration of oxygenated and deoxygenated hemoglobin. One of the key features of fNIRS is its superior temporal resolution, with dense measurements over very short periods of time (100ms increments). Unfortunately, most statistical analysis approaches in the existing literature have not fully utilized the high temporal resolution of fNIRS. For example, many analysis procedures are based on linearity assumptions that only extract partial information, thereby neglecting the overall dynamic trends in fNIRS trajectories. The main goal of this article is to assess the ability of a functional data analysis approach for detecting significant differences in hemodynamic responses recorded by fNIRS. Children with and without specific language impairment wore two, 3×5 fNIRS caps situated over the bilateral parasylvian areas as they completed a language comprehension task. Functional data analysis was used to decompose the high dimensional hemodynamic curves into the mean function and a few eigenfunctions to represent the overall trend and variation

structures over time. Compared to the most popular general linear model, we did not assume any parametric structure and let the data speak for itself. This analysis identified significant differences between the case and control groups in the oxygenated hemodynamic mean trends in the right *inferior frontal cortex* and left *inferior posterior parietal cortex* brain regions. We also detected significant group differences in the deoxygenated hemodynamic mean trends in the right *inferior posterior parietal cortex* and left *temporal parietal junction* brain region. These findings, using dramatically different approaches, experimental designs, data sets, and foci, were consistent with several other reports, confirming group differences in the importance of these two areas for syntax comprehension. The proposed functional data analysis was consistent with the temporal characteristics of fNIRS, thus providing an alternative methodology for fNIRS analyses.

Keywords: fNIRS, Hemodynamic response curve, Functional Data Analysis, Specific Language Impairment, Sentence Comprehension

1 INTRODUCTION

Functional near infrared spectroscopy (fNIRS) is a non-invasive method for measuring near-infrared light absorption through the skull, enabling researchers to speculate a close proxy to neural activation that results from relative changes of the cerebrovascular alterations in oxygenated and deoxygenated hemoglobin concentrations in cortical structures (Villringer and Dirnagl, 1994; Boas et al., 2014; Tak and Ye, 2014). Since light between 650 and 950 nm is weakly absorbed by biological chromophores (Hoge et al., 2005), the relatively deep penetration of NIR light makes it an effective research tool in neuro-imaging studies. Compared to other imaging technologies such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), fNIRS has a few advantages such as low cost, high flexibility, portability, and the ability to accommodate young children and patients with psychological issues (Arenth et al., 2007; Ye et al., 2009). fNIRS offers superior temporal resolution with dense measurements over time and provides data for a wide range of functional contrasts such as oxygenated (ΔHbO), deoxygenated (ΔHbD), and total hemoglobin (ΔHbT) responses simultaneously as participants perform functional tasks in naturalistic environments (Ye et al., 2009; Kozel et al., 2009; Tak and Ye, 2014; Hall et al., 2013). Despite the extensive study of fNIRS data, little has been done to study the mean and variation trends of hemodynamic curves as individuals complete language processing tasks. Indeed, analysis approaches that truly utilize the superior temporal characteristics of fNIRS are rare in the existing literature. Even rarer are studies of concomitant behavioral and neural differences between children with specific language impairment (SLI) and typically developing control children as they complete language comprehension tasks.

In this article, we introduce a functional data analysis (FDA) methodology with a goal of addressing several challenging questions: 1) how to best utilize the superior temporal resolution of fNIRS; 2) how to model its hemodynamic trends for syntax-related stimuli; 3) how to connect light optodes with brain regions without anatomy information; 4) how to speculate the differences in brain activities between case and control in reaction to the same stimuli. FDA is a nonparametric data-driven statistical technique that does not make any parametric assumption such as the linearity or normality. Our main objective was to model the overall hemodynamic trends from a functional perspective as opposed to individual discrete points that are considered using existing analysis approaches. Although the modeling goal of FDA conforms to the temporal hemodynamic signals of the fNIRS context (Barati et al., 2013), it has seldom been applied in the fNIRS literature.

Tak and Ye (2014) reviewed currently existing statistical models in fNIRS data. The most well-known and widely used method was the general linear model (GLM) (Schroeter et al., 2004; Plichta et al., 2007), which has been integrated into numerous fNIRS analysis tools (Shimada and Hiraki, 2006; Koh et al., 2007; Abdelnour and Huppert, 2009; Ye et al., 2009; Strangman et al., 2009; Huppert et al., 2009; Custo et al., 2010; Penny et al., 2011). As a multivariate statistical model, GLM works well, but FDA differs in important ways. First, GLM is a traditional parametric model that assumes a linear combination structure. Assuming a parametric form would likely be misleading if the underlying data did not satisfy the main linear assumptions. Therefore, nonparametric modeling without any assumptions should be more flexible. Second, as a multivariate model, GLM does not utilize the time course of the data and hence can not capture the overall trends of the hemoglobin concentration in the dynamic or functional sense (Barati et al., 2013). Third, GLM does not provide a relevant hypothesis test approach to compare the differences in the overall hemodynamic trends between case and control groups due to its model structure restrictions.

Comparing which brain regions are significantly involved in a task performed by two groups requires formal hypothesis testing. Unfortunately, many of the current statistical approaches used to perform hypothesis tests for fNIRS data may not be optimal in the functional sense. Simple statistics such as t-test have been performed to statistically compare single-value differences between different groups (Germon et al., 1994; Aldrich et al., 1994; Germon et al., 1999; Young et al., 2000; Hoshi et al., 2001; Isobe et al., 2001; Kennan et al., 2002; Schroeter et al., 2002; Hoshi, 2003; Matsuo et al., 2003; Tachtsidis et al., 2004; Tsujimoto et al., 2004; Shibuya-Tayoshi et al., 2007; Kim et al., 2010). Multi-way ANOVA has also been employed in fNIRS studies (Fallgatter and Strik, 1998; Bartocci et al., 2000; Fallgatter and Strik, 2000; Herrmann et al., 2003; Hoshi, 2003; Suto et al., 2004; Folley and Park, 2005; Kameyama et al., 2006; Arenth et al., 2007; Irani et al., 2007). Although these methods were able to evaluate differences in hemoglobin observations, information was lost because only partial measurements were considered. Using FDA to compare the overall temporal mean and variation trends of hemodynamic functions rather than simply defining a magnitude may be more informative and robust, especially in a context in which optical signal attenuation or motion artifacts cause noise (Ye et al., 2009).

When repeated measurements are recorded over a dense grid of time points, often by machine, they are typically termed as functional or longitudinal data, with one observed curve per subject. Formally, FDA models each hemodynamic response curve as a continuum function over time, thus capturing the overall dynamic trajectories of the function over time, even though the measurements are collected discretely (Ramsay and Silverman, 2002; Ramsay, 2006; Ferraty and Vieu, 2006; Barati et al., 2013). Although some experimental errors are generally unavoidable, the nonparametric kernel smoothing captures the underlying mean function and hence greatly reduces the effects of noise. The functional principal component analysis (FPCA) based on the Karhunen-Loeve theorems decomposes the high dimensional auto-covariance matrix extracted from fNIRS data to a few important orthogonal eigenfunctions. The first few eigenfunctions explaining the majority of variation are likely induced by cognitive related tasks, with the remaining eigenfunctions explaining only a very small percentage of variation that may be caused by nuisance factors such as breathing, vasomotor, measurement error, movement artifacts, and other unaccounted activities (Akgül et al., 2006). To perform comprehensive comparisons on the hemodynamic curves between case and control groups, we tested the equality of mean functions, and eigenfunctions and eigenvalues of the auto-covariance functions using two-sample FPCA approaches. Bootstrap sampling was used to determine the threshold of the significance of the tests because the distributions of the test statistics were unknown (Benko et al., 2009). Importantly, FDA is inherently nonparametric and does not assume any parametric structure or distributions within the hemodynamic curve data.

Some researchers have investigated the functional relationship between fNIRS and fMRI and their correlation over time (Mandeville et al., 1999; Siegel et al., 2003; Okamoto et al., 2004; Fujiwara et al., 2004; Steinbrink et al., 2006). Although many common properties exist between fNIRS and fMRI, functional curve based modeling, which is mature in fMRI research (Grodzinsky, 2000; Ben Schachar et al., 2003; Müller et al., 2003; Ben-Shachar et al., 2004; Binder et al., 2009; Seghier et al., 2010; Seghier, 2013; Weismer et al., 2005), has rarely been used for fNIRS stand-alone experiments. The progress achieved in fMRI analyses paves the way for improvements on fNIRS approaches. We believe that the FDA approach could promote breakthroughs in fNIRS research, similar to the way it did for fMRI.

To test the potential of FDA to analyze fNIRS data, we used fNIRS to asssess differences in neural activation between children (case: children with specific language impairment; control: age-matched, typically-developing children) as they engaged a language comprehension task that is known to favor the children in the control group. Specific language impairment is a developmental language disorder of unknown origin that is characterized by significant deficits in the acquisition and use of spoken and written language in the absence of hearing, intellectual, emotional, or acquired neurological impairments (Leonard, 2014; Bishop, 2014). This disorder affects approximately 7% of the school-age population (Tomblin et al., 1997). If functional data analysis is a promising statistical approach for fNIRS, it should reveal group differences in parasylvian (language related) neural regions as children perform the task.

2 MATERIALS AND METHODS

2.1 Participants

Thirty children (15 children with specific language impairment and 15 age-matched, typically developing control children) between the ages of 8 and 12 participated in the study. There were 8 males in each group. The children in the SLI group met the standard classification criteria of performance on multiple language measures that was one or more standard deviations below the mean. The typically-developing controls performed above one standard deviation from the mean on multiple language measures. All the children in both groups were right-handed, monolingual English speakers. All the children in the SLI group were receiving special education services in the public schools. In addition, we provided independent testing to insure that the children in the SLI group met our identification criteria.

2.2 Sentence Comprehension Task

The children completed a language comprehension task in which they listened to a sentence and then selected a picture (from three choices) that depicted the agent (actor) in the sentence. There were 60 total sentences with 15 sentences representing each of four sentence types: subject-verb-object (“The ring had moved the square behind the very bright cold bed”), subject relatives (“The watch that had hugged the truck behind the kite was bright”), passives (“The shoe was hugged by the clock under the very cold box”), and object relatives (“The book that the shirt had hugged under the kite was new”). The sentences were controlled for length, vocabulary complexity, and vocabulary imageability (Montgomery et al., 2015). Similar to Dick et al. (2004), noun animacy and noun affordance cues were removed, making the sentences semantically implausible. This was done so that the children’s decisions about the agent of the sentence would be based primarily on syntactic knowledge or word order rather than semantic plausibility. Children saw three pictures on a computer screen as they listened to each sentence. They were asked to point to the picture of the agent of the sentence (the thing doing the action) as quickly as possible after hearing each

sentence. All children completed 8 training items before fNIRS scanning began. See Montgomery et al. (2015) for a complete description of the stimuli.

2.3 functional Near Infrared Spectroscopy Procedures

Data was collected with the Hitachi ETG-4000 (Hitachi Medical Co., Japan) with 44 channels divided across two 3×5 probe caps. The channels were determined by bilateral placement of the optode caps such that the middle detector in the lowest row of optodes was placed over T3 or T4. The measurement patches covered the majority of the right and left parasylvian regions including *inferior frontal cortex*, *inferior parietal lobule* (including the *temporal parietal junction* and *inferior posterior parietal cortex*), and superior temporal cortex. The channel locations are depicted in Figure 1.

The fNIRS scan began with a 30 second rest period in which children were instructed to focus on a "+" in the middle of the computer screen and to "relax" their mind. After the first rest period, children listened to 60, 12-word sentences representing four different syntax types (15 subject-verb-object sentences, 15 subject relative clause sentences, 15 passive sentences, and 15 object relative clause sentences). E-prime software was used to present the stimuli in a pseudo-random order and to record the accuracy and speed of the children's responses. The sentences were presented in three blocks of 20 items, presented in a psudorandom order, with each item being separated by a jittered rest interval that varied between 2 and 6 seconds. Each block was separated by a 25 second rest period. The stimuli onsets for each participant were consistently predefined and each participant was given 8 seconds to think and respond.

Throughout the fNIRS scan, near-infrared light from the source optodes travels approximately 1-1.5 cm into the cortex where it is absorbed by oxygen molecules attached to hemoglobin in the blood in the brain (Dehghani and Delpy, 2000). The amount of light that is not absorbed is measured by the detecting optodes. The relative changes in the concentration of oxygenated hemoglobin (ΔHbO), deoxygenated hemoglobin (ΔHbD) and total hemoglobin (ΔHbT) were estimated according to changes in the optical properties of the light using the Beers-Lambert conversion (see Plichta et al. (2007) for a detailed description). A length of 8,521 and a frequency of 10 Hz time series was collected within a duration of 851 seconds for each channel of each participant. Figure 2A displays one example of original ΔHbO time series at channel 31 (mainly overlapped in the right *inferior frontal cortex*) for a child in the SLI group.

2.4 Data Preprocessing

There were a total of 3,960 individual time series collected from three hemoglobin categories (ΔHbO, ΔHbD and ΔHbT), 44 channels, and 30 participants (15 cases and 15 controls). Each time series contained 8,521 measurement units consisting of 4,800 intermittent task measurement units and 3,721 rest measurement units. The active periods represented 15 stimuli segments for each of the four syntax types. The following preprocessing steps were designed to extract the most important information from such a large amount of data.

The first step of data preprocessing was to group channels based on regions of interest (ROIs). The global alignments of the channel positions between individuals were difficult because fNIRS has the shortcoming of weak spatial anatomical representation. The ROIs for the current project were derived a priori based on previous findings in both the fMRI and fNIRS literature demonstrating changes in cortical activation during language processing tasks. Four areas within the parasylvian region, *inferior frontal cortex* (*Broca's* area), *superior temporal cortex*, the *temporal parietal junction* and *posterior inferior parietal cortex* (*Angular Gyrus*) are frequently implicated in verbal tasks (Rossi et al., 2012; Scherer et al., 2012; Petrides, 2013). A Polemus system was used for 3D digitization of head size and optode location following testing. This

provided standardized Montreal Neurological Institute coordinates and anatomical labels that related to each participant individually. We determined the corresponding channel for each monitored brain region based on the largest percentage of overlapping rate between the channel and the brain regions of interest for each participant.

The second step of data preprocessing was to extract only stimulus-related active units from the original time series and focus only on the segments associated with cognitive activity during the target stimulus comprehension tasks. There were 60 such windowed segments, each lasting 8 seconds (corresponding to 80 units), and hence, a total 4,800 units were extracted. As an example, Figure 2B displays the stimulus-relevant ΔHbO extracted from the original time series at channel 31 (mainly overlapping the right *inferior frontal cortex*) for a child (Sue) in the SLI group. This process was repeated for all individuals.

Since the observations were collected very densely, we used the average of the 10 units per second as the modeling target, illustrated in Figure 2C. Comparing B and C of Figure 2, notice that the two signals look almost the same, except Figure 2B has length 4,800 but Figure 2C is only of length 480 (1/10 of original length). If there were any differences caused by averaging the 10 dense units per second (Figure 2C), it would be smoother and would capture the trend even better by removing more noise or errors from averaging.

The third step of data preprocessing related to selecting the hemoglobin categories. It is not clear whether neuronal activation is best represented by ΔHbO, ΔHbD, or ΔHbT. Researchers may expect that the deoxygenated hemoglobin to show opposite trends to that of ΔHbO because the ΔHbO and ΔHbD often complement each other (Cui et al., 2010). However, comparing Figure 3A with Figure 3B for one example of the same channel for the same person, note that the deoxygenated hemodynamic trends are flatter than the oxygenated hemodynamic trends, and there does not appear to be opposite trends in most time segments. This suggests that the oxygenated hemoglobin contains a more rubust signal than the deoxygenated hemoglobin. In this article, we mainly focused on modeling ΔHbO and ΔHbD because the results of ΔHbT (the sum of ΔHbO and ΔHbD) were highly correlated with the other two.

The forth step of data preprocessing involved extracting the syntax-relevant time course by locating the time onsets of the 15 questions for each syntax type. This yielded four different time courses, each with 120 units. As an example, Figure 4 displays one example of the four syntax-relevant ΔHbO hemodynamic curves extracted from the original time series at channel 31 (mainly overlapped in the right *inferior frontal cortex*) for a child in the SLI group named Sue.

Factors such as breathing, vasomotor response, measurement error, movement artifacts, and other unaccounted activities (Akgül et al., 2006), may cause noise in fNIRS data. These four preprocessing steps enabled us to extract the most important signals and remove unavoidable confounding factors. Comparing Figure 2A with Figure 4, notice that it is harder to recognize patterns from Figure 2A due to many complex and sharp fluctuations and strands. On the contrary, the patterns are smoother and clearer in Figure 4. After these preprocessing steps, our data were ready for the statistical models and hypothesis tests.

2.5 Functional Data Analysis Structure

Let $Y_{ikc}, i = 1, \ldots, n;\ k = 1, \ldots, T;\ c = 1, 2$, denote the relative changes in the concentrations of oxygenated or deoxygenated hemoglobin of the ith subject measured at discrete time point t_k for the cth group. Here $c = 1$ denotes the case group and $c = 2$ denotes the control group, n is the number of subjects per group, and T is the total time points measured for each subject. These observed densely collected curves with noise can be modeled as independent realizations of a stochastic process with smooth trajectories.

Let $X_{1c}(t), \ldots, X_{nc}(t)$ denote random smooth trajectories of the underlying stochastic process in $L^2(\mathcal{T})$, $t \in \mathcal{T}$, where $\mathcal{T}$ is the time interval. Then we can reconstruct the smooth functions $X_i's$ from the original densely collected noisy observations $Y_i's$ as (Müller, 2008)

$$Y_{ikc} = X_{ic}(t) + \varepsilon_{ikc},\ i = 1, \ldots, n;\ c = 1, 2;\ k = 1, \ldots, T;\ t \in \mathcal{T}, \tag{1}$$

where ε_{ikc} are the experimental errors and assumed to be independent, with $E(\varepsilon_{ikc}) = 0$ and $Var(\varepsilon_{ikc}) = \sigma^2_{kc}$.

For each group c, the mean function of $X_{ic}(t)$ is $\mu_c(t) = E(X_{ic}(t))$ and auto-covariance function of $X_{ic}(t)$ is

$$G_c(s,t) = cov\{X_{ic}(s), X_{ic}(t)\} = E\{[X_{ic}(s) - \mu_c(s)][X_{ic}(t) - \mu_c(t)]\},$$

for $s, t \in \mathcal{T}$. Here $\mu_c(t)$ is interpreted as the mean function of oxygenated or deoxygenated hemodynamic curves for group c. Throughout this paper, it is assumed that $\mu_c(t)$ is a smooth function of t, and $G_c(s,t)$ is a positive definite and bivariate smooth function of s and t, for $s, t \in \mathcal{T}$. The "smooth" refers to twice continuously differentiable. The idea of model (1) is that the observed noisy curve over time is described by an underlying smooth function plus noise.

In order to model the auto-covariance function, functional PCA interprets $G_c(s,t)$ as the kernel of a linear integral operator on the space $L_2(\mathcal{T})$ of square-integrable functions on $\mathcal{T}$, mapping $f \in L_2(\mathcal{T})$ to $A_{G_c} f \in L_2(\mathcal{T})$ defined by

$$(A_{G_c} f)(t) = \int_{\mathcal{T}} f(s) G_c(s,t) ds. \tag{2}$$

An eigenfunction v of the auto-covariance operator A_{G_c} is a solution of the equation $(A_{G_c} v)(t) = \lambda v(t)$, with eigenvalue λ. For each c, we assume that the operator A_{G_c} has a sequence of smooth orthonormal eigenfunctions v_{lc} satisfying $\int_{\mathcal{T}} v_{kc}(t) v_{lc}(t) dt = \delta_{kl}$ (here δ_{kl} is the Kronecker symbol), with ordered eigenvalues $\lambda_{1c} \geq \lambda_{2c} \geq \ldots \geq 0$. By Mercer's Theorem, applying a spectral decomposition on the function G_c yields

$$G_c(s,t) = \sum_{l=1}^{\infty} \lambda_{lc} v_{lc}(s) v_{lc}(t). \tag{3}$$

Since the eigenfunctions v_{lc}'s form a complete orthonormal sequence on $L_2(\mathcal{T})$, the generalized Fourier expansion ($Karhunen - Loeve$ Theorem (Karhunen, 1946) or functional principal component expansion) on X_{ic} yields

$$X_{ic}(t) = \mu_c(t) + \sum_{l=1}^{\infty} \zeta_{ilc} v_{lc}(t), c = 1, 2, \tag{4}$$

where the sum is defined in the sense of L_2 convergence and

$$\zeta_{ilc} =< X_{ic} - \mu_c, v_{lc} >= \int_{\mathcal{T}} (X_{ic}(t) - \mu_c(t)) v_{lc}(t) dt \tag{5}$$

are uncorrelated random variables with $E(\zeta_{ilc}) = 0$, and $var(\zeta_{ilc}) = \lambda_{lc}$, subject to the L_2 convergence, i.e.

$$\Sigma_l \lambda_{lc} = E(||X_{ic} - \mu_c||^2) = \int G_c(t,t) dt < \infty.$$

ζ_{lc} are frequently referred to as the lth functional principal component score or the lth dominant modes of random effects.

By way of Equation (4), the dynamic trends of random function $X_{ic}(t)$ can be modeled by the mean trend function $\mu_c(t)$, the eigenfunction v_{lc}, and the distribution of functional principal component scores ζ_{ilc}. The first L principal components were used to approximate Equation (4) to capture the most important variations, remove the noise effects, and estimate the main signals of the trajectories of $X_c(t)$ effectively (Ramsay and Silverman, 2002).

2.6 Parameter Estimates

Using the observed data set $\mathcal{D} = \{Y_{ikc}, i = 1, \ldots, n;\ k = 1, \ldots, T;\ c = 1, 2\}$, we were able to estimate all unknown parameters $\hat{\mu}_c(t)$, $\hat{G}_c(s,t)$, and $\hat{\sigma}^2_{kc}$ from Equations (1 - 5). The smooth function $X_{ic}(t_k)$ and $\hat{\sigma}^2_{kc}$ of each discrete noisy observation (t_{ik}, Y_{ikc}) were estimated by model (1) via nonparametric kernel smoothing. Then the unbiased estimator of $\mu_c(t)$ was easily obtained from the sample mean of $X_{ic}(t)$.

Once the estimator $\hat{\mu}_c(t)$ was obtained, we computed the sample estimate of auto-covariance matrix by

$$\hat{G}_c(t,s) = n^{-1}\Sigma_{i=1}^{n}\{X_{ic}(s) - \hat{\mu}_c(s)\}\{X_{ic}(t) - \hat{\mu}_c(t)\}.$$

The estimate of eigenfunctions were obtained by the corresponding spectral decomposition on $\hat{G}_c(s,t)$. To be more specific, $\hat{\lambda}_{qc}$ are eigenvalues of $\hat{G}_c$, given by

$$\int_{\mathcal{T}} \hat{G}_c(s,t)\hat{v}_{lc}(s)ds = \hat{\lambda}_{lc}\hat{v}_{lc}(t).$$

And $\hat{v}_{lc}$ is the eigenfunction corresponding to $\hat{\lambda}_{lc}$, satisfying $\int_{\mathcal{T}} \hat{v}^2_{lc}(t)dt = 1$ and $\int_{\mathcal{T}} \hat{v}_{kc}\hat{v}_{lc}(t)dt = 0$ if $k \neq l$. The signs of $\hat{v}_{lc}$ were not uniquely determined. In order to ensure the closeness of $\hat{v}_{lc}$ from two groups of $c = 1, 2$, we allowed the signs of $\hat{v}_{lc}$ to be chosen arbitrarily as long as $< \hat{v}_{l1}, \hat{v}_{l2} > \ \geq 0$ for $l = 1, \ldots, L$.

$\hat{G}_c$ also presents an empirical version of the expansion (3)

$$\hat{G}_c(s,t) = \sum_{l=1}^{L} I(\hat{\lambda}_{lc} > 0)\hat{\lambda}_{lc}\hat{v}_{lc}(s)\hat{v}_{lc}(t), \tag{6}$$

where I is the indicator function used to only keep the terms with positive eigenvalues. From the percentage of variation explained by the first few eigenfunctions, the first L largest eigenvalues $\hat{\lambda}_{1c}, \ldots, \hat{\lambda}_{Lc}$ were chosen. The positive definiteness of the estimated auto-covariance matrix $\hat{G}_c(s,t)$ was not always guaranteed, which might be a problem in practical applications. Once $\hat{\lambda}_{lc}$ and $\hat{v}_{lc}$ were obtained, we checked whether or not $\hat{\lambda}_{lc} > 0$ (Müller, 2008). If $\hat{\lambda}_{lc}$ was negative, then we dropped this negative eigenvalue and its corresponding eigenfunction, and reconstituted the estimate from remaining eigenvalues and eigenfunction estimates.

Once eigenvalues $\hat{\lambda}_{1c} \geq \ldots \geq \hat{\lambda}_{Lc}$ and orthonormal eigenfunctions $\hat{v}_1, \ldots, \hat{v}_L$ were obtained, the fitting of individual trajectories required estimation of functional principal component scores. By the discretization on the equation (5), plugging $\hat{\mu}_c$ and $\hat{v}_{lc}$ into a Riemann sum approximation of the integral, we have

$$\hat{\zeta}_{ilc} = \Sigma_{k=1}^{T}(X_{ic}(t_k) - \hat{\mu}(t_k))\hat{v}_{lc}(t_k)(t_k - t_{k-1}), \tag{7}$$

setting $t_0 = 0$ (Müller, 2008). We assured that $n^{-1}\Sigma\hat{\zeta}_{ilc} = 0$, $n^{-1}\Sigma\hat{\zeta}_{ilc}\hat{\zeta}_{iwc} = 0$ for $l \neq w$; $l, w = 1, \ldots, L$, and $n^{-1}\Sigma\hat{\zeta}^2_{ilc} = \hat{\lambda}_{lc}$. This approximation method by sum worked well because our observations were collected densely and consistently for all subjects without missing values.

2.7 Nonparametric Kernel Smoothing

The nonparametric regression kernel smoothing was a traditional approach to capture the curve trends without making assumptions about the error distributions. The goal of smoothing was to model the underlying function by estimating $X(t) = E(Y|t)$ from the original discrete measurement and removing the noisy observations caused by measurement errors. To define a kernel smoother, we need a bandwidth h and a kernel function K.

The Nadaraya-Watson Estimator (NW), a basic framework for kernel estimators (Nadaraya, 1964; Watson, 1964; Cai, 2001; Racine and Li, 2004; Bailey et al., 2010; Demir and Toktamiş, 2010; Kato, 2012; Simonoff, 2012), was defined by

$$\frac{\sum_{i=1}^{n} K_h(t - t_i)Y_i}{\sum_{j=1}^{n} K_h(t - t_j)}, \tag{8}$$

where $K_h(t) = 1/hK(t/h)$. The kernel function $K(t)$ was a non-negative symmetric real valued integrable function satisfying $\int_{-\infty}^{\infty} K(t)dt = 1, \int_{-\infty}^{\infty} tK(t)dt = 0$, and $\int_{-\infty}^{\infty} t^2K(t)dt > 0$. The Epanechnikov kernel $K(t) = 3/4(1 - t^2)I(|t| < 1)$ was used. The bandwidth h controled the number of points that neighbored each t_i and hence determined the weight of each point contributing to the estimator. The choice of bandwidth was crucial in changing the result because it served as a smoothing parameter and determined the trade-off between the variance and bias of the resulting nonparametric regression estimates. Typically, smaller h decreases the bias but increases the estimation variance. We chose the optimal bandwidth that minimized the Generalized Cross Validation (GCV).

$$GCV(h) = \frac{1}{T(1 - \nu/T)^2} \sum_{k=1}^{T} (Y_{ikc} - X_{ic}(t_k))^2,$$

for each subject i and group c. Here ν is the trace of matrix M

$$M = \begin{pmatrix} l_1(t_1) & l_2(t_2) & \ldots & l_T(t_1) \\ l_1(t_2) & l_2(t_2) & \ldots & l_T(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ l_1(t_T) & l_2(t_T) & \ldots & l_T(t_T) \end{pmatrix},$$

with

$$l_i(t) = \frac{K_h(t - t_i)}{\sum_{j=1}^{n} K_h(t - t_j)}$$

Once the smooth trajectory of each $X_i, i = 1, \ldots, n$ was estimated from the NW nonparametric kernel smoother with the optimal bandwidth, we estimated the mean $\hat{\mu}_c(t)$ for each group directly from the sample mean, which was a consistent and unbiased estimator.

2.8 Hypothesis Tests

The main goal of this article was to determine whether functional data analysis applied to fNIRS data would reveal significant differences in the hemodynamic function curves between the case and control

groups as they processed syntax-related stimuli. We examined eight parasylvian brain regions: left and right *inferior frontal cortex*, the *temporal parietal junction*, *inferior posterior parietal cortex*, and *superior temporal cortex*. Statistically, we used formal hypothesis tests to judge the extent to which the distributions of the random functions $X_{1c}, \ldots, X_{nc}$ differed for case and control groups. By way of the empirical Karhunen-Loeve decompositions (4), we approximated the functions of $X_{ic}(t)$ as

$$X_{ic}(t) = \hat{\mu}_c(t) + \sum_{l=1}^{L} \hat{\zeta}_{ilc}\hat{v}_{lc}(t), c = 1, 2;\ i = 1, \ldots, n. \tag{9}$$

As a result, the possible differences of the hemodynamic signals between the case and control group could be tested from the following three steps.

The first test was whether or not significant differences existed for the overall mean trends between case and control group for each syntax type at each brain area of interest:

$$H_{01} : \mu_1(t) = \mu_2(t), t \in \mathcal{T}.$$

If H_{01} failed to be rejected, it would mean that the overall mean trends of hemodynamic curves were similar between the case and control groups. The second test was whether or not significant differences existed for the variation trends between case and control groups for each syntax type at each brain area of interest:

$$H_{0,2l} : v_{l1}(t) = v_{l2}(t),\ t \in \mathcal{T}; l = 1, \ldots, L.$$

If $H_{0,2l}$ failed to be rejected, it would mean that the l^{th} variation mode had similar trends between the case and control groups. The third test was whether or not significant differences existed for the variance of principal component scores for each syntax type at each brain area of interest:

$$H_{0,3l} : \lambda_{l1} = \lambda_{l2},\ l = 1, \ldots, L.$$

If $H_{0,3l}$ failed to be rejected, it would mean that distribution of the l^{th} principal component scores were similar between the case and control group.

The first two tests, H_{01} and $H_{0,2l}$ were challenging because they were based on high dimensional curves, and both the test statistics and the distribution were unknown. The most traditional approach involves judging the similarity of two curves by measuring how far the norm of the differences of the two vectors is away from zero. Define the following measures (Benko et al., 2009):

$$\begin{aligned} D_1 &= ||\hat{\mu}_1(t) - \hat{\mu}_2(t)||^2, \\ D_{2,l} &= ||\hat{v}_{l1}(t) - \hat{v}_{l2}(t)||^2, l = 1, \ldots, L, \\ D_{3,l} &= |\hat{\lambda}_{l1} - \hat{\lambda}_{l2}|^2, l = 1, \ldots, L. \end{aligned}$$

The three null-hypotheses would be rejected respectively, if

$$D_1 \geq \Delta_{1;1-\alpha};\ \ D_{2,l} \geq \Delta_{2,l;1-\alpha};\ \ D_{3,l} \geq \Delta_{3,l;1-\alpha},$$

where $\Delta_{1;1-\alpha}$, $\Delta_{2,l;1-\alpha}$, and $\Delta_{3,l;1-\alpha}$ denotes the α-level critical values of the distributions of

$$\Delta_1 = ||(\hat{\mu}_1(t) - \mu_1(t)) - (\hat{\mu}_2(t) - \mu_2(t))||^2,$$
$$\Delta_{2,l} = ||(\hat{v}_{l1}(t) - v_{l1}(t)) - (\hat{v}_{l2}(t) - v_{l2}(t))||^2, \ \ l = 1, \ldots, L,$$
$$\Delta_{3,l} = |(\hat{\lambda}_{l1} - \lambda_{l1}) - (\hat{\lambda}_{l2} - \lambda_{l2})|^2, \ \ l = 1, \ldots, L.$$

We decided to use Δs as the primary test because Ds were equal to Δs under the null hypotheses and the values of Ds were shifted by the difference in the true means, eigenfunctions, and eigenvalues under the alternative hypotheses. However, because the true population mean, eigenvalues and eigenfunctions were unknown, above Δs can not be accessed directly. Therefore, we used the bootstrap sampling to determine the threshold (Benko et al., 2009).

$$\Delta_1^* = ||(\hat{\mu}_1(t) - \hat{\mu}_1^*(t)) - (\hat{\mu}_2(t) - \hat{\mu}_2^*(t))||^2,$$
$$\Delta_{2,l}^* = ||(\hat{v}_{l1}(t) - \hat{v}_{l1}^*(t)) - (\hat{v}_{l2}(t) - \hat{v}_{l2}^*(t))||^2, \ \ l = 1, \ldots, L,$$
$$\Delta_{3,l}^* = |(\hat{\lambda}_{l1} - \hat{\lambda}_{l1}^*) - (\hat{\lambda}_{l2} - \hat{\lambda}_{l2}^*)|^2, \ \ l = 1, \ldots, L,$$

where $\hat{\mu}_1^*(t), \hat{v}_{l1}^*(t), \hat{\lambda}_{l1}^*(t)$, as well as $\hat{\mu}_2^*(t), \hat{v}_{l2}^*(t), \hat{\lambda}_{l2}^*(t)$ were estimated from each independent bootstrap samples $X_{11}^*(t), \ldots, X_{n1}^*(t)$ and $X_{12}^*(t), \ldots, X_{n2}^*(t)$, respectively. We performed 1,000 nonparametric bootstrap samples for both case and control group and we repeated the nonparametric kernel smoothing for each sample. Finally the $1 - \alpha$ percentiles were used to determine the thresholds of the tests.

3 RESULTS

3.1 Real NIRS Data Analysis

Behaviorally, the case (specific language impairment) group identified the agents of subject-verb-object and subject relative clause sentences as well as their age-matched, typically developing controls. However, the children in the case group were significantly less accurate than the children in the control group on the passive and object relative clause sentences.

The goal of statistical modeling was to determine whether there were significant differences in the hemodynamic trends between the case and control groups. Additionally, we speculated which brain regions were associated with children's syntax comprehension ability from the significant group differences. For each hemodynamic category (ΔHbO and ΔHbD), we performed 32 tests to consider all combinations of four different syntax types and eight different brain regions.

Using the functional data analysis approaches described in Sections 2.5 and 2.6, we first estimated the mean function $\hat{\mu}_c(t)$, eigenfunctions $\hat{v}_{lc}(t)$, and eigenvalues $\hat{\lambda}_{lc}$ for each group, with $c = 1$ corresponding to case group and $c = 2$ for control group. During the analysis, we kept the first two eigenfunctions (i.e. L=2) because they explained 90% of the overall variations, and the remaining eigenfunctions explained only a very small percentage of the variations.

With respect to potential group differences in mean trends of ΔHbO, H_{01} was rejected at the significance level of 0.1 at two brain regions: right *inferior frontal cortex* brain region for subject-verb-object, subject relative clause, and object relative clause sentences, and at the left *inferior posterior parietal cortex* brain region for object relative clause and passive sentences. Therefore, we concluded that the right

Table 1 Significant group differences in the percentages of variation explained by the first two eigenfunctions between the case and control groups (Note: IPPC stands for *inferior posterior parietal cortex*; TPJ stands for *temporal parietal junction*; STC stands for *superior temporal cortex*; OR stands for object relative clause sentences; SR for subject relative clause sentences; SVO for subject-verb-object sentences; and PAS for passive sentences; v_{11} stands for the first eigenfunction of the case group; v_{21} stands for the second eigenfunction of the case group; v_{12} stands for the first eigenfunction of the control group; and v_{22} stands for the second eigenfunction of the control group).

Category	Brain Region	Syntax	Case		Control	
			v_{11}	v_{21}	v_{12}	v_{22}
ΔHbO	left IPPC	OR	88.3%	5.5%	76.8%	15.3%
ΔHbO	left TPJ	PAS	97.4%	0.6%	87.9%	5.2%
ΔHbO	left STC	PAS	89.8%	3.9%	92.8%	5.4%
ΔHbO	left STC	SVO	84.7%	6.0%	95.2%	1.8%
ΔHbO	right STC	OR	83.0%	6.2%	96.1%	1.3%
ΔHbD	left IPPC	OR	97.5%	0.9%	59.8%	33.6%
ΔHbD	left IPPC	PAS	97.6%	0.8%	61.2%	31.4%
ΔHbD	left IPPC	SR	97.7%	0.7%	66.8%	23.8%
ΔHbD	left IPPC	SVO	96.8%	1.0%	69.8%	15.6%
ΔHbD	left TPJ	OR	97.1%	0.8%	83.0%	12.8%
ΔHbD	left TPJ	PAS	97.1%	0.8%	87.6%	8.3%
ΔHbD	left TPJ	SR	97.9%	0.6%	91.8%	4.5%
ΔHbD	left STC	OR	93.8%	2.4%	83.5%	13.3%
ΔHbD	left STC	PAS	92.1%	2.7%	87.7%	10.0%
ΔHbD	left STC	SR	95.0%	1.9%	88.8%	8.1%
ΔHbD	left STC	SVO	92.4%	2.6%	89.2%	5.5%

inferior frontal cortex and left *inferior posterior parietal cortex* were associated with the children's syntax comprehension processing ability. Figure 5 displays the estimated mean trajectories $\hat{\mu}_c(t)$ of ΔHbO in these two brain regions with corresponding significant syntax types. A close inspection of Figure 5 reveals that the mean trajectories of case and control have different dynamic trends (different shape and magnitude) for each syntax type, with opposite fluctuate oscillations at some time segments but similar directions at other time segments. The mean trajectories of the control group were always above those of the case group in these two brain regions. In the right *inferior frontal cortex* brain region, the mean oxygenated hemodynamic trajectories of the control group were always above zero, while those of the case group were below zero. In the left *inferior posterior partietal cortex*, the mean oxygenated hemodynamic trajectories of both case and control groups were below zero.

The hypothesis test H_{01} for ΔHbD was rejected at the right *inferior posterior parietal cortex* (at 0.05 significance level) and left *temporal parietal junction* (at 0.1 significance level) for all four syntax types. We concluded that there were significant differences (related to both shape and magnitude) in the mean trajectories of ΔHbD at these two brain regions between case and control group, and these two brain regions were also associated with children's syntax comprehension ability. A close inspection of Figure 6 reveals that the mean trajectories of ΔHbD for the control group mainly fluctuate around zero but that of case group around -0.2 for all the eight scenarios. Using the same range of y-axis as the oxygenated hemodynamic trajectories of ΔHbO in Figure 5, the overall mean trends of the deoxygenated hemodynamic trajectories ΔHbD were very flat, especially those of the case group. So, we decreased the range of the y-axis in Figure 6 to the half of that of Figure 5 so that the significant oscillations were more apparent.

None of the 32 hypothesis tests related to the variation trends ($H_{0,2l}, l = 1, 2$) could be rejected for either ΔHbO or ΔHbD at any of the eight brain regions or for any of the four syntax type types. Thus, there were no significant differences in the eigenfunction (i.e. variation trends) in ΔHbD and ΔHbD between the case and control group.

Hypothesis test $H_{0,3l}, l = 1, 2$, related to the eigenvalues, was rejected at a few brain regions and syntax types. It indicated that the percentages of variation explained by the first two eigenfunctions (i.e. the distributions of the first two principle component scores) were significantly different between case and control groups. Table 1 summarizes the details of percentage of variation for all significant brain regions and syntax types. Among all these significant differences, the left *inferior posterior parietal cortex* brain region for ΔHbD achieved the maximum for all four syntax types, with the first eigenfunction of the case group ($v_{11}(t)$) explaining $96 - 98\%$ of the total variation of the case group versus $59 - 70\%$ of the total variation of the control group ($v_{12}(t)$). Similarly, the second eigenfunction ($v_{21}(t)$) explained $0.7 - 1.0\%$ of the total variation of the case group versus $15 - 34\%$ of the total variation of the control group ($v_{22}(t)$). Additionally, we also noticed that the *superior temporal cortex* brain regions for ΔHbO showed opposite directions in the percentage of variation explained by the first two eigenfuncitons as compared to other brain regions. Specifically, the first eigenfunction of the case group ($v_{11}(t)$) explained a greater percentage of total variation than the first eigenfunction of the control group ($v_{12}(t)$) for almost all scenarios, except the ΔHbO at left *superior temporal cortex* for passive and subject-verb-objects sentences, and right *superior temporal cortex* for object relative clause sentences. Also, we observed that the second eigenfunction of the case group ($v_{21}(t)$) explained a much smaller percentage of total variation than the second eigenfunction of the control group ($v_{22}(t)$) for almost all scenarios with the exception of the ΔHbO at left *superior temporal cortex* for subject-verb-object sentences and right *superior temporal cortex* for object relative clause sentences.

4 DISCUSSION

The primary goal of this article was to determine whether significant group differences in the hemodynamic trajectories existed for two groups with known language differences. To achieve this goal, we designed a syntax type comprehension tasks in which 15 children with specific language impairments and 15 age-matched, typically-developing controls pointed to pictures representing the agent (actor) after hearing four types of sentences (subject-verb-object sentences, subject relative clause sentences, passive sentences, and object relative clause sentences). We administered the 60 questions in a pseudo-random order to 30 participants during the NIRS data collection. We performed three formal hypothesis tests to formally assess the group differences between the case and control group, and determined the threshold by bootstrap

approach for high dimensional object when both test statistics and distributions were unknown (Benko et al., 2009).

The functional data analysis approach is different from the widly used traditional approaches in existing NIRS literature (e.g., GLM and t-test). In functional data analyis, the modeling is performed in the functional sense that treats the entire curve as the modeling target and fully utilizes the superior temporal resolution of fNIRS data. But GLM extracts multivariate discrete points and does not utilize the dynamic trajectories of the fNIRS curve. As a nonparametric data-driven approach, FDA does not assume any linear structure or normality distribution such as that within the GLM model (Shimada and Hiraki, 2006; Koh et al., 2007; Abdelnour and Huppert, 2009; Custo et al., 2010; Penny et al., 2011; Tak and Ye, 2014). Unlike simple t tests (Germon et al., 1994; Aldrich et al., 1994; Germon et al., 1999; Young et al., 2000; Hoshi et al., 2001; Isobe et al., 2001; Kennan et al., 2002; Schroeter et al., 2002; Hoshi, 2003; Matsuo et al., 2003; Tachtsidis et al., 2004; Tsujimoto et al., 2004; Shibuya-Tayoshi et al., 2007; Kim et al., 2010), FDA tests the trajectory differences of two entire curves for two groups and captures not only the differences in magnitude but also in shape. Thus, our approach was inclusive of all observed stimulus-relevant data information and was not restricted to the magnitudee differences as t-test does.

We successfully detected significant group differences in the oxygenated hemodynamic mean trends in two brain regions, right *inferior frontal cortex* and left *inferior posterior parietal cortex*. The mean oxygenated hemodynamic trajectories between case and control groups showed different trends (different shape and magnitude) in these two brain regions, with some segments showing opposite fluctuating oscillations but other segments having similar directions. In the right *inferior frontal cortex* brain region, the mean oxygenated hemodynamic trajectories of the control group were always above zero, while those of the case group were below zero. In the left *inferior posterior partietal cortex*, the mean oxygenated hemodynamic trajectories of both case and control groups were below zero. We also detected significant group differences in deoxygenated hemodynamic mean trends in the region of the right *inferior posterior partietal cortex* and left *temporal parietal junction* brain area. The mean deoxygenated hemodynamic trajectories of the control group mainly fluctuated around the zero line while that of case group were all below -0.2. Some of these significant findings from our quantitative functional NIRS analysis were consistent with the results of a few other studies that had dramatically different approaches, experiments, data sets, and foci. For example, the left *inferior posterior parietal cortex (Angular Gyrus)* brain region has been reported to be highly engaged in semantic processing during language comprehension (Geschwind, 1965; Joseph, 1982; Demonet et al., 1992; Vandenberghe et al., 1996; Vigneau et al., 2006; Houdé et al., 2010; Price, 2010), including some reports got by MRI (Binder et al., 2009; Seghier et al., 2010; Seghier, 2013). Further, differences between children with and without SLI in the extent of activation of this area has been noted in studies of listening to nonwords and words (Weismer et al., 2005). A number of MRI studies have noted group differences between children with SLI and their age-matched controls in the size of right hemisphere parasylvian areas (Plante et al., 1991).

There were no significant differences in the eigenfunctions, but the percentage of total variation explained by each eigenfunction significantly differenced in the left *inferior posterior partietal cortex*, left *temporal parietal junction*, and both left and right *superior temporal cortex*. The finding of significant group differences in the percentage of variation explained by the first two eigenfunctions may be of particular interest. Recall that the first two orthogonal eigenfunctions derived from the fNIRS high dimensional auto covariance matrix were likely related to the cognitive processes involved in performing our syntax comprehension task. The significant group differences in the percentage of total variation explained by the eigenfunctions may relate to group differences in information processing functions that have been

associated with attention, semantic processing, and syntactic processing in the left *inferior posterior parietal cortex*, the left *temporal parietal junction*, and the left *superior temporal cortex*. Further research on larger samples of participants are needed to fully understand the meaning of these results.

In future work, we will compare the signicant differences between left and right hemispheres. Unlike the comparisons between case and control groups, the left and right brain samples are not independent requiring a different approach. We will also explore more detailed functional properties in the rest periods. Although there are several hypothesis tests involved, we will leave the multiple correction for the future for a few reasons. First, there are only 15 subjects within each group, which is much less than the dimension of the curves (length of 120 after preprocessing and length of 8,521 before preprocessing). As a result, power is limited due to the difficulties of collecting children with SLI. Therefore, we do not want to diminish our ndings due to a large number of multiple corrections. We believe that our methods will yield better results after the sample size is large enough and will investigate the multiple correction when we have an appropriate sample size. Second, the multiple tests involved here are not independent. Instead, they form close correlations, as ΔHbO and ΔHbD and the four syntax types are highly correlated. Therefore, many multiple correction approaches will not be appropriate and likely will mislead the results. For example, we nd that the test of equal mean hemodynamic trends between case and control ($H_{1,0}$) reject, whether we consider each of the syntax types (with 120 length) individually or we test the stimuli of the four syntax types simultaneously (with 480 length). However, if we use multiple correction, say Bonferroni correction, then each syntax test will only have an $\alpha/4$ signicance level, which makes the individual syntax period impossible to be rejected given the current sample size. In that case, none of the individual syntax types would show signicant differences, but the whole stimuli curve with four syntax types will be signicant between case and control. This will result in conicting conclusions.

In summary, this proof of concept study was conducted to explore a more advanced statistical analysis approach to the analysis of the time course of hemodynamic data collected with functional near infrared spectroscopy. This approach enables us to compare which brain regions are signicantly involved in syntax comprehension ability in the two groups. Functional data analysis strategies were used to decompose the high dimensional ΔHbO and ΔHbD time curves into mean curves and eigenfunctions to represent overall trends and variation structures (Ramsay and Silverman, 2002; Ramsay, 2006; Ferraty and Vieu, 2006; Barati et al., 2013). After detailed comparisons and hypothesis tests, we revealed greater brain activity for the case group than the control group for all four syntax types. In addition, different percentages of variation for the case and control groups were explained by the first two eigenfunctions, suggesting that the two groups used different cognitive processing strategies while performing the tasks. The approach of FDA proposed in this paper has promise as an analysis method that captures the overall mean trends and variation trends of hemoglobin concentration over time within and between groups without assuming any structure.

ETHICS STATEMENT

This study was approved by the Utah State University Institutional Review Board. All participants (adults and children) and the parents or guardians of all children signed consent forms that were approved by the IRB. The participant's name "Sue" was a pseudonym.

DISCLOSURE/CONFLICT-OF-INTEREST STATEMENT

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

GF conceived the statistical modeling, programmed and performed the data analysis and figures, and wrote the first version of the manuscript; NW collected and processed the data; JMB assisted with experimental design and programmed the task; JWM and JLE created the experimental task; RBG conceived the research, helped design the experimental task, supervised all aspects of data collection and data processing, edited the manuscript, and wrote the experimental design section.

Funding: Funded in part by the Lillywhite Endowment to Utah State University.

REFERENCES

Abdelnour, A. F. and Huppert, T. (2009). Real-time imaging of human brain function by near-infrared spectroscopy using an adaptive general linear model. *Neuroimage* 46, 133–143

Akgül, C. B., Akin, A., and Sankur, B. (2006). Extraction of cognitive activity-related waveforms from functional near-infrared spectroscopy signals. *Medical and Biological Engineering and Computing* 44, 945–958

Aldrich, C., Wyatt, J., Spencer, J., Reynolds, E., and Delpy, D. (1994). The effect of maternal oxygen administration on human fetal cerebral oxygenation measured during labour by near infrared spectroscopy. *BJOG: An International Journal of Obstetrics & Gynaecology* 101, 509–513

Arenth, P. M., Ricker, J. H., and Schultheis, M. T. (2007). Applications of functional near-infrared spectroscopy (fnirs) to neurorehabilitation of cognitive disabilities. *The Clinical Neuropsychologist* 21, 38–57

Bailey, R. W., Addison, J. T., et al. (2010). A smoothed-distribution form of nadaraya-watson estimation. *Department of Economics Discussion Paper* , 10–30

Barati, Z., Zakeri, I., and Pourrezaei, K. (2013). Functional data analysis view of functional near infrared spectroscopy data. *Journal of biomedical optics* 18, 117007–117007

Bartocci, M., Winberg, J., Ruggiero, C., Bergqvist, L. L., Serra, G., and Lagercrantz, H. (2000). Activation of olfactory cortex in newborn infants after odor stimulation: a functional near-infrared spectroscopy study. *Pediatric Research* 48, 18–23

Ben Schachar, M., Hendler, T., Kahn, I., Ben Bashat, D., and Grodzinsky, Y. (2003). The neural reality of syntactic transformations. *Psychological Science* 14, 433–440

Ben-Shachar, M., Palti, D., and Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fmri experiments. *Neuroimage* 21, 1320–1336

Benko, M., Härdle, W., Kneip, A., et al. (2009). Common functional principal components. *The Annals of Statistics* 37, 1–34

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19, 2767–2796

Bishop, D. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders* 49, 381–415

Boas, D. A., Elwell, C. E., Ferrari, M., and Taga, G. (2014). Twenty years of functional near-infrared spectroscopy: introduction for the special issue. *NeuroImage* 85, 1–5

Cai, Z. (2001). Weighted nadaraya–watson regression estimation. *Statistics & probability letters* 51, 307–318

Cui, X., Bray, S., and Reiss, A. L. (2010). Functional near infrared spectroscopy (nirs) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* 49, 3039–3046

Custo, A., Boas, D. A., Tsuzuki, D., Dan, I., Mesquita, R., Fischl, B., et al. (2010). Anatomical atlas-guided diffuse optical tomography of brain activation. *Neuroimage* 49, 561–567

Dehghani, H. and Delpy, D. T. (2000). Near-infrared spectroscopy of the adult head: effect of scattering and absorbing obstructions in the cerebrospinal fluid layer on light distribution in the tissue. *Applied optics* 39, 4721–4729

Demir, S. and Toktamiş, Ö. (2010). On the adaptive nadaraya-watson kernel regression estimators. *Hacettepe Journal of Mathematics and Statistics* 39

Demonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768

Dick, F., Wulfeck, B., Krupa-Kwiatkowski, M., and Bates, E. (2004). The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science* 7, 360–377

Fallgatter, A. J. and Strik, W. K. (1998). Frontal brain activation during the wisconsin card sorting test assessed with two-channel near-infrared spectroscopy. *European archives of psychiatry and clinical neuroscience* 248, 245–249

Fallgatter, A. J. and Strik, W. K. (2000). Reduced frontal functional asymmetry in schizophrenia during a cued continuous performance test assessed with near-infrared spectroscopy. *Schizophrenia bulletin* 26, 913–919

Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice* (Springer Science & Business Media)

Folley, B. S. and Park, S. (2005). Verbal creativity and schizotypal personality in relation to prefrontal hemispheric laterality: A behavioral and near-infrared optical imaging study. *Schizophrenia research* 80, 271–282

Fujiwara, N., Sakatani, K., Katayama, Y., Murata, Y., Hoshino, T., Fukaya, C., et al. (2004). Evoked-cerebral blood oxygenation changes in false-negative activations in bold contrast functional mri of patients with brain tumors. *Neuroimage* 21, 1464–1471

Germon, T., Evans, P., Barnett, N., Wall, P., Manara, A., and Nelson, R. (1999). Cerebral near infrared spectroscopy: emitter-detector separation must be increased. *British journal of anaesthesia* 82, 831–837

Germon, T., Kane, N., Manara, A., and Nelson, R. (1994). Near-infrared spectroscopy in adults: effects of extracranial ischaemia and intracranial hypoxia on estimation of cerebral oxygenation. *British journal of anaesthesia* 73, 503–506

Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain* 88, 585–585

Grodzinsky, Y. (2000). The neurology of syntax: Language use without broca's area. *Behavioral and brain sciences* 23, 1–21

Hall, M., Chaudhary, U., Rey, G., and Godavarty, A. (2013). Fronto-temporal mapping and connectivity using nirs for language-related paradigms. *Journal of Neurolinguistics* 26, 178–194

Herrmann, M., Ehlis, A.-C., and Fallgatter, A. (2003). Frontal activation during a verbal-fluency task as measured by near-infrared spectroscopy. *Brain Research Bulletin* 61, 51–56

Hoge, R., Franceschini, M., Covolan, R., Huppert, T., Mandeville, J., and Boas, D. (2005). Simultaneous recording of task-induced changes in blood oxygenation, volume, and flow using diffuse optical imaging and arterial spin-labeling mri. *Neuroimage* 25, 701–707

574 Hoshi, Y. (2003). Functional near-infrared optical imaging: Utility and limitations in human brain mapping.
575 *Psychophysiology* 40, 511–520
576 Hoshi, Y., Kobayashi, N., and Tamura, M. (2001). Interpretation of near-infrared spectroscopy signals: a
577 study with a newly developed perfused rat brain model. *Journal of Applied Physiology* 90, 1657–1662
578 Houdé, O., Rossi, S., Lubin, A., and Joliot, M. (2010). Mapping numerical processing, reading, and
579 executive functions in the developing brain: an fmri meta-analysis of 52 studies including 842 children.
580 *Developmental science* 13, 876–885
581 Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). Homer: a review of
582 time-series analysis methods for near-infrared spectroscopy of the brain. *Applied optics* 48, 280–298
583 Irani, F., Platek, S. M., Bunce, S., Ruocco, A. C., and Chute, D. (2007). Functional near infrared
584 spectroscopy (fnirs): an emerging neuroimaging technology with important applications for the study of
585 brain disorders. *The Clinical Neuropsychologist* 21, 9–37
586 Isobe, K., Kusaka, T., Nagano, K., Okubo, K., Yasuda, S., Kondo, M., et al. (2001). Functional imaging
587 of the brain in sedated newborn infants using near infrared topography during passive knee movement.
588 *Neuroscience letters* 299, 221–224
589 Joseph, R. (1982). The neuropsychology of development: Hemispheric laterality, limbic language, and the
590 origin of thought. *Journal of clinical psychology* 38, 4–33
591 Kameyama, M., Fukuda, M., Yamagishi, Y., Sato, T., Uehara, T., Ito, M., et al. (2006). Frontal lobe
592 function in bipolar disorder: a multichannel near-infrared spectroscopy study. *Neuroimage* 29, 172–184
593 Karhunen, K. (1946). Zur spektraltheorie stochastischer prozesse
594 Kato, K. (2012). Weighted nadaraya–watson estimation of conditional expected shortfall. *Journal of*
595 *Financial Econometrics* 10, 265–291
596 Kennan, R. P., Kim, D., Maki, A., Koizumi, H., and Constable, R. T. (2002). Non-invasive assessment
597 of language lateralization by transcranial near infrared optical topography and functional mri. *Human*
598 *brain mapping* 16, 183–189
599 Kim, M. N., Durduran, T., Frangos, S., Edlow, B. L., Buckley, E. M., Moss, H. E., et al. (2010). Noninvasive
600 measurement of cerebral blood flow and blood oxygenation using near-infrared and diffuse correlation
601 spectroscopies in critically brain-injured adults. *Neurocritical care* 12, 173–180
602 Koh, P. H., Glaser, D. E., Flandin, G., Kiebel, S., Butterworth, B., Maki, A., et al. (2007). Functional
603 optical signal analysis: a software tool for near-infrared spectroscopy data processing incorporating
604 statistical parametric mapping. *Journal of biomedical optics* 12, 064010–064010
605 Kozel, F. A., Tian, F., Dhamne, S., Croarkin, P. E., McClintock, S. M., Elliott, A., et al. (2009).
606 Using simultaneous repetitive transcranial magnetic stimulation/functional near infrared spectroscopy
607 (rtms/fnirs) to measure brain activation and connectivity. *Neuroimage* 47, 1177–1184
608 Leonard, L. B. (2014). *Children with specific language impairment* (MIT press)
609 Mandeville, J. B., Marota, J. J., Ayata, C., Moskowitz, M. A., Weisskoff, R. M., and Rosen, B. R. (1999).
610 Mri measurement of the temporal evolution of relative cmro 2 during rat forepaw stimulation. *Magnetic*
611 *resonance in medicine* 42, 944–951
612 Matsuo, K., Taneichi, K., Matsumoto, A., Ohtani, T., Yamasue, H., Sakano, Y., et al. (2003). Hypoactivation
613 of the prefrontal cortex during verbal fluency test in ptsd: a near-infrared spectroscopy study. *Psychiatry*
614 *Research: Neuroimaging* 124, 1–10
615 Montgomery, J. W., Evans, J. L., Gillam, R. B., Sergeev, A. V., and Finney, M. C. (2015). "whatdunit?"
616 developmental changes in children's syntactically-based sentence interpretation abilities and sensitivity
617 to word order (in press). *Applied Psycholinguistics* 1, 1–12
618 Müller, H.-G. (2008). Functional modeling of longitudinal data. *Longitudinal Data Analysis* 1, 223–252

Müller, R.-A., Kleinhans, N., and Courchesne, E. (2003). Linguistic theory and neuroimaging evidence: An fmri study of brocas area in lexical semantics. *Neuropsychologia* 41, 1199–1207

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications* 9, 141–142

Okamoto, M., Dan, H., Shimizu, K., Takeo, K., Amita, T., Oda, I., et al. (2004). Multimodal assessment of cortical activation during apple peeling by nirs and fmri. *Neuroimage* 21, 1275–1288

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images* (Academic press)

Petrides, M. (2013). *Neuroanatomy of language regions of the human brain* (Academic Press)

Plante, E., Swisher, L., Vance, R., and Rapcsak, S. (1991). Mri findings in boys with specific language impairment. *Brain and language* 41, 52–66

Plichta, M., Heinzel, S., Ehlis, A.-C., Pauli, P., and Fallgatter, A. (2007). Model-based analysis of rapid event-related functional near-infrared spectroscopy (nirs) data: a parametric validation study. *Neuroimage* 35, 625–634

Price, C. J. (2010). The anatomy of language: a review of 100 fmri studies published in 2009. *Annals of the New York Academy of Sciences* 1191, 62–88

Racine, J. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* 119, 99–130

Ramsay, J. O. (2006). *Functional data analysis* (Wiley Online Library)

Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis: methods and case studies*, vol. 77 (Springer New York)

Rossi, S., Telkemeyer, S., Wartenburger, I., and Obrig, H. (2012). Shedding light on words and sentences: near-infrared spectroscopy in language research. *Brain and language* 121, 152–163

Scherer, L. C., Fonseca, R. P., Amiri, M., Adrover-Roig, D., Marcotte, K., Giroux, F., et al. (2012). Syntactic processing in bilinguals: An fnirs study. *Brain and language* 121, 144–151

Schroeter, M. L., Bücheler, M. M., Müller, K., Uludağ, K., Obrig, H., Lohmann, G., et al. (2004). Towards a standard analysis for functional near-infrared imaging. *NeuroImage* 21, 283–290

Schroeter, M. L., Zysset, S., Kupka, T., Kruggel, F., and Von Cramon, D. Y. (2002). Near-infrared spectroscopy can detect brain activity during a color–word matching stroop task in an event-related design. *Human brain mapping* 17, 61–71

Seghier, M. L. (2013). The angular gyrus multiple functions and multiple subdivisions. *The Neuroscientist* 19, 43–61

Seghier, M. L., Fagan, E., and Price, C. J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *The Journal of Neuroscience* 30, 16809–16817

Shibuya-Tayoshi, S., Sumitani, S., Kikuchi, K., Tanaka, T., Tayoshi, S., UENO, S.-I., et al. (2007). Activation of the prefrontal cortex during the trail-making test detected with multichannel near-infrared spectroscopy. *Psychiatry and clinical neurosciences* 61, 616–621

Shimada, S. and Hiraki, K. (2006). Infant's brain responses to live and televised action. *Neuroimage* 32, 930–939

Siegel, A. M., Culver, J. P., Mandeville, J. B., and Boas, D. A. (2003). Temporal comparison of functional brain imaging with diffuse optical tomography and fmri during rat forepaw stimulation. *Physics in medicine and biology* 48, 1391

Simonoff, J. S. (2012). *Smoothing methods in statistics* (Springer Science & Business Media)

Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the bold signal: combined fmri–fnirs studies. *Magnetic resonance imaging* 24, 495–505

Strangman, G. E., Zhang, Q., and Zeffiro, T. (2009). Near-infrared neuroimaging with ninpy. *Frontiers in Neuroinformatics* 3

Suto, T., Fukuda, M., Ito, M., Uehara, T., and Mikuni, M. (2004). Multichannel near-infrared spectroscopy in depression and schizophrenia: cognitive brain activation study. *Biological Psychiatry* 55, 501–511

Tachtsidis, I., Elwell, C. E., Leung, T. S., Lee, C.-W., Smith, M., and Delpy, D. T. (2004). Investigation of cerebral haemodynamics by near-infrared spectroscopy in young healthy volunteers reveals posture-dependent spontaneous oscillations. *Physiological measurement* 25, 437

Tak, S. and Ye, J. C. (2014). Statistical analysis of fnirs data: a comprehensive review. *NeuroImage* 85, 72–91

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., and O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research* 40, 1245–1260

Tsujimoto, S., Yamamoto, T., Kawaguchi, H., Koizumi, H., and Sawaguchi, T. (2004). Prefrontal cortical activation associated with working memory in adults and preschool children: an event-related optical topography study. *Cerebral cortex* 14, 703–712

Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiak, R. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature* 383, 254–6

Vigneau, M., Beaucousin, V., Herve, P.-Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432

Villringer, A. and Dirnagl, U. (1994). Coupling of brain activity and cerebral blood flow: basis of functional neuroimaging. *Cerebrovascular and brain metabolism reviews* 7, 240–276

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A* , 359–372

Weismer, S. E., Plante, E., Jones, M., and Tomblin, J. B. (2005). A functional magnetic resonance imaging investigation of verbal working memory in adolescents with specific language impairment. *Journal of Speech, Language, and Hearing Research* 48, 405–425

Ye, J. C., Tak, S., Jang, K. E., Jung, J., and Jang, J. (2009). Nirs-spm: statistical parametric mapping for near-infrared spectroscopy. *Neuroimage* 44, 428–447

Young, A., Germon, T., Barnett, N., Manara, A., and Nelson, R. (2000). Behaviour of near-infrared light in the adult human head: implications for clinical near-infrared spectroscopy. *British Journal of Anaesthesia* 84, 38–42

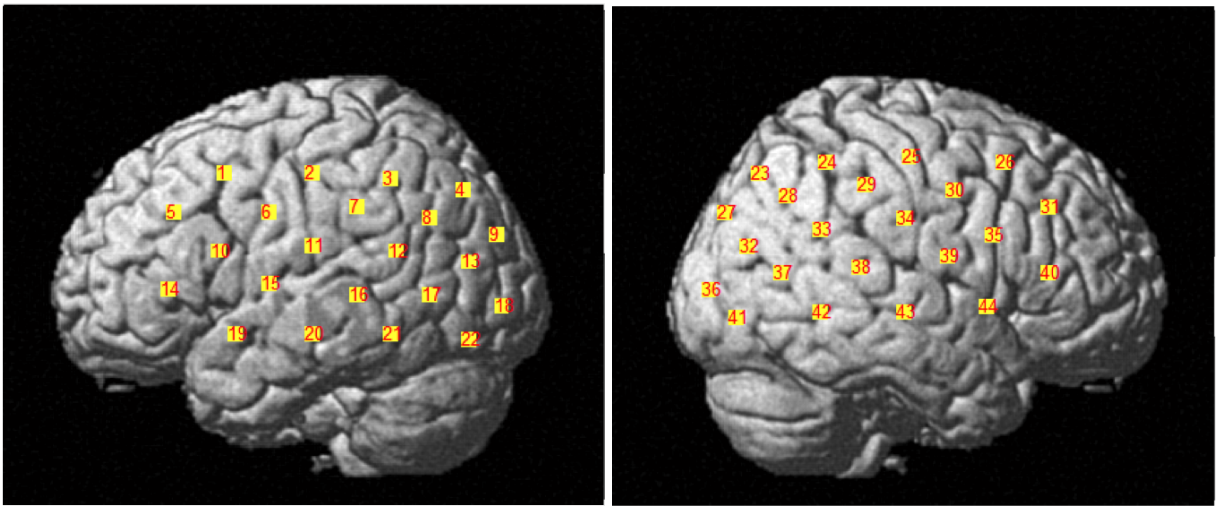

Figure 1: Display of the 44 channels divided across two 3 $\times$ 5 probe caps. The channels 1-22 belong to the left brain hemisphere and the channels 23-44 belong to the right brain hemisphere.

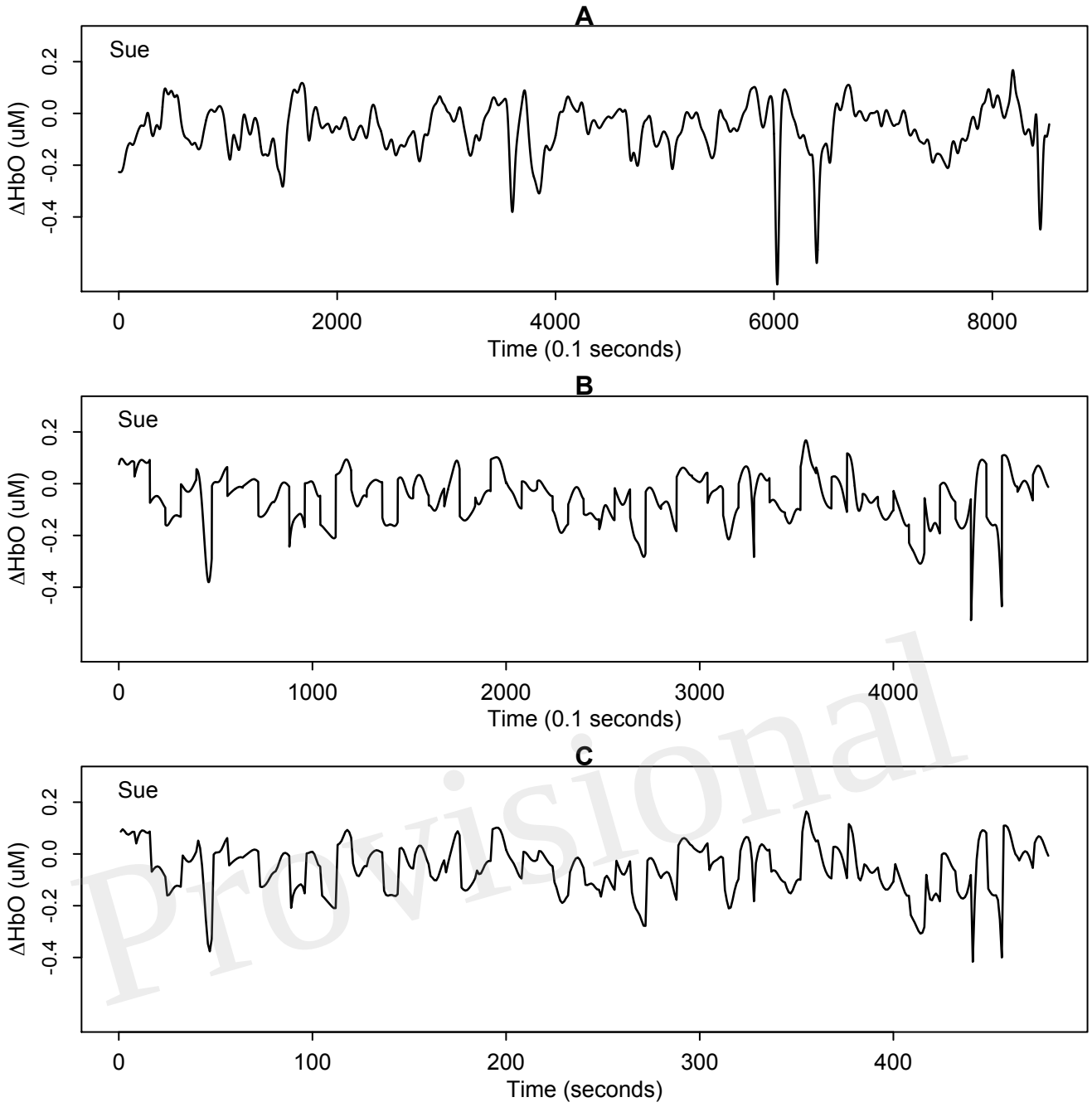

Figure 2: One example of ΔHbO time series at channel 31 (mainly overlapped in the right *inferior frontal cortex*) for Sue, a child participant with specific language impairment. A: the original time series of ΔHbO with length 8,521; B: the extracted stimulus-relevant ΔHbO under 60 target stimuli instants with length 4,800; C: the average version of B with length 480. By averaging the 10 measurements of each second, the curve maintains similar signal but only using 1/10 of original length.

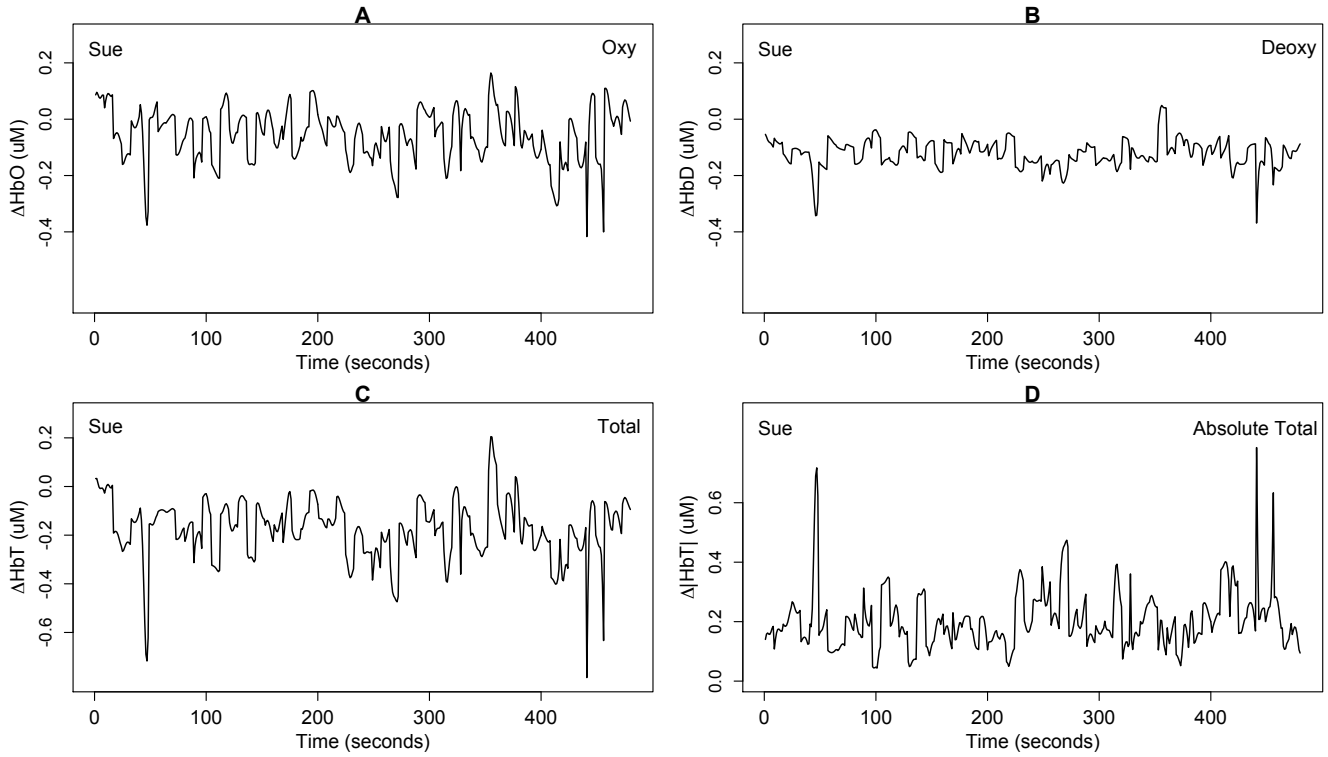

Figure 3: One example of the four different stimulus-relevant hemoglobin categories at channel 31 (mainly overlapped in the right *inferior frontal cortex*) for Sue. A: the stimulus-relevant oxygenated hemodynamic curve ΔHbO; B: the stimulus-relevant deoxygenated hemodynamic curve ΔHbD; C: the stimulus-relevant total hemodynamic curve ΔHbT; D: the stimulus-relevant absolute total hemodynamic curve $\Delta|HbT|$. ΔHbT is computed by summing ΔHbO and ΔHbD. Absolute total $\Delta|HbT|$ was computed by summing the absolute value of ΔHbO and ΔHbD.

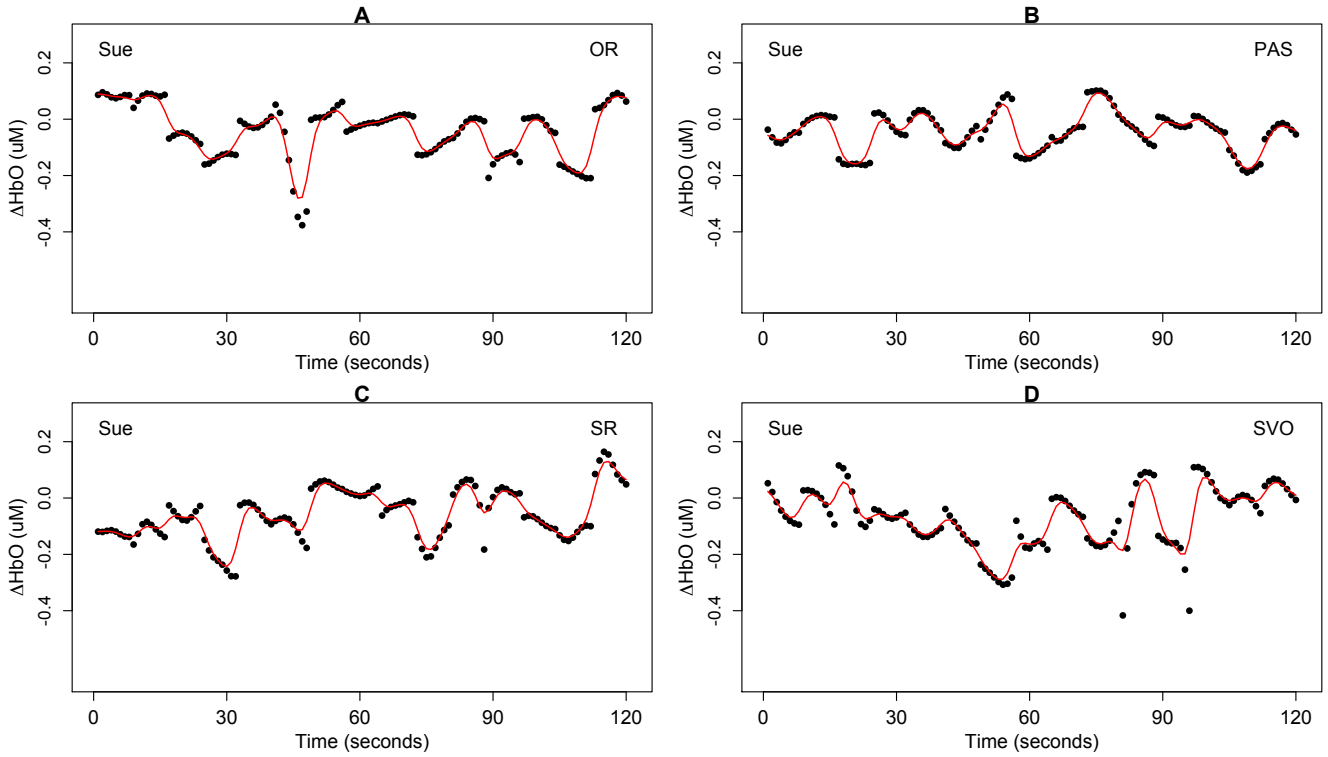

Figure 4: One example of extracted syntax-relevant ΔHbO at channel 31 (mainly overlapped in the right *inferior frontal cortex*) for Sue under four syntax types respectively, each with 15 target stimulus questions. The black dots are the original observation Y_{ikc} of oxygenated hemoglobin and the red curves are smoothing hemodynamic trajectories $X_{ic}(t)$ estimated by nonparametric kernel smoother from model (1). A: object relative clause sentences (OR); B: passive sentences (PAS); C: subject relative clause sentences (SR); D: subject-verb-object sentences (SVO).

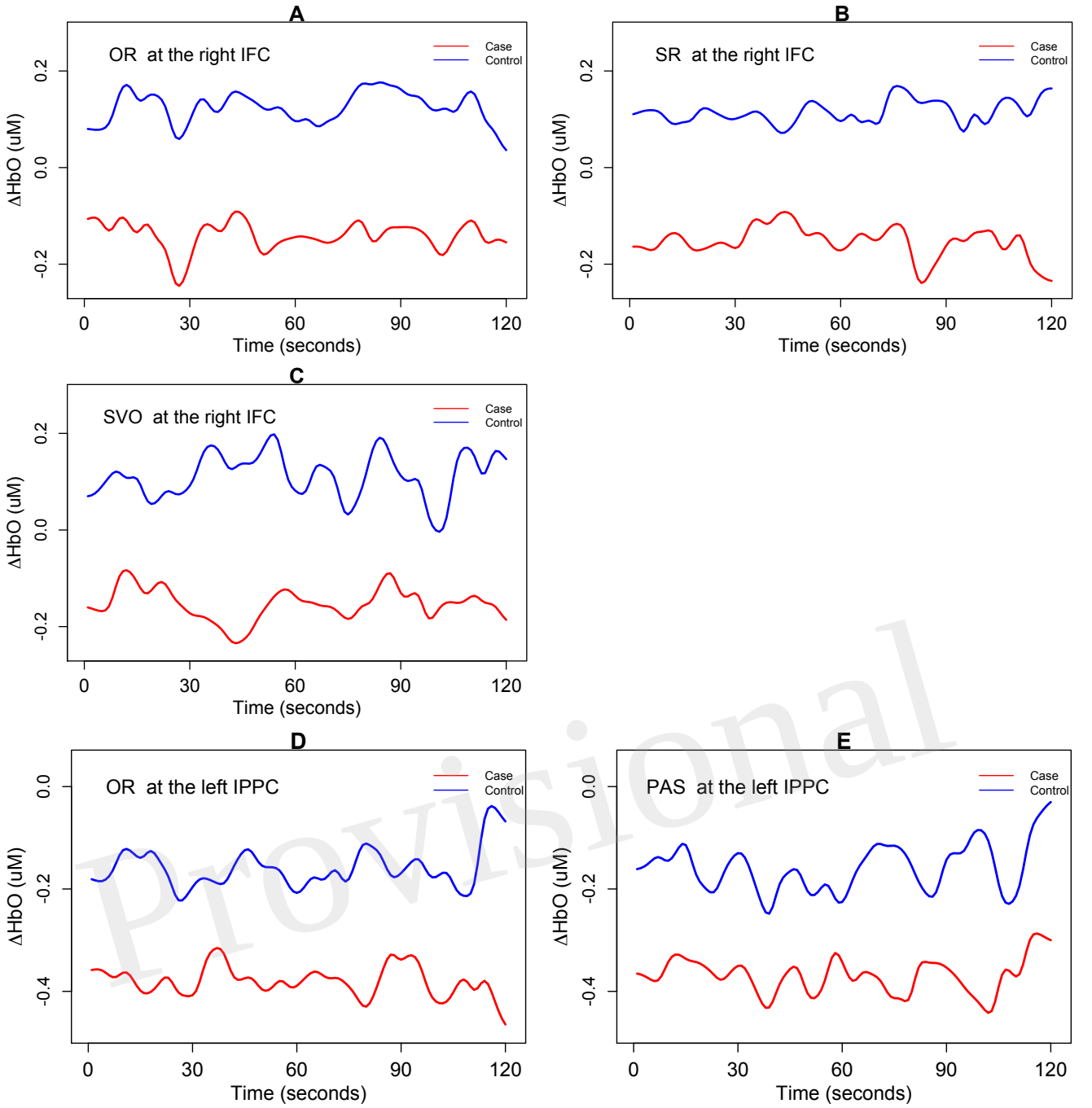

Figure 5: The mean trends for ΔHbO (i.e. H_{01}). The mean trajectories $\hat{\mu}_c(t)$ of 15 smoothing ΔHbO curves, in the control group ($c = 2$) is depicted as a blue line and similar information for case group ($c = 1$) is depicted as a red line. IFC = *inferior frontal cortex*; IPPC = *inferior posterior parietal cortex*. OR = object relative clause sentences; SR = subject relative clause sentences; SVO = subject-verb-object sentences; and PAS = passive sentences. A: the mean trajectories of ΔHbO for OR syntax type at the right IFC; B: the mean trajectories of ΔHbO for SR syntax type at the right IFC; C: the mean trajectories of ΔHbO for SVO syntax type at the right IFC; D: The mean trajectories of ΔHbO for OR syntax type at the left IPPC; E: the mean trajectories of ΔHbO for PAS syntax type at the left IPPC.

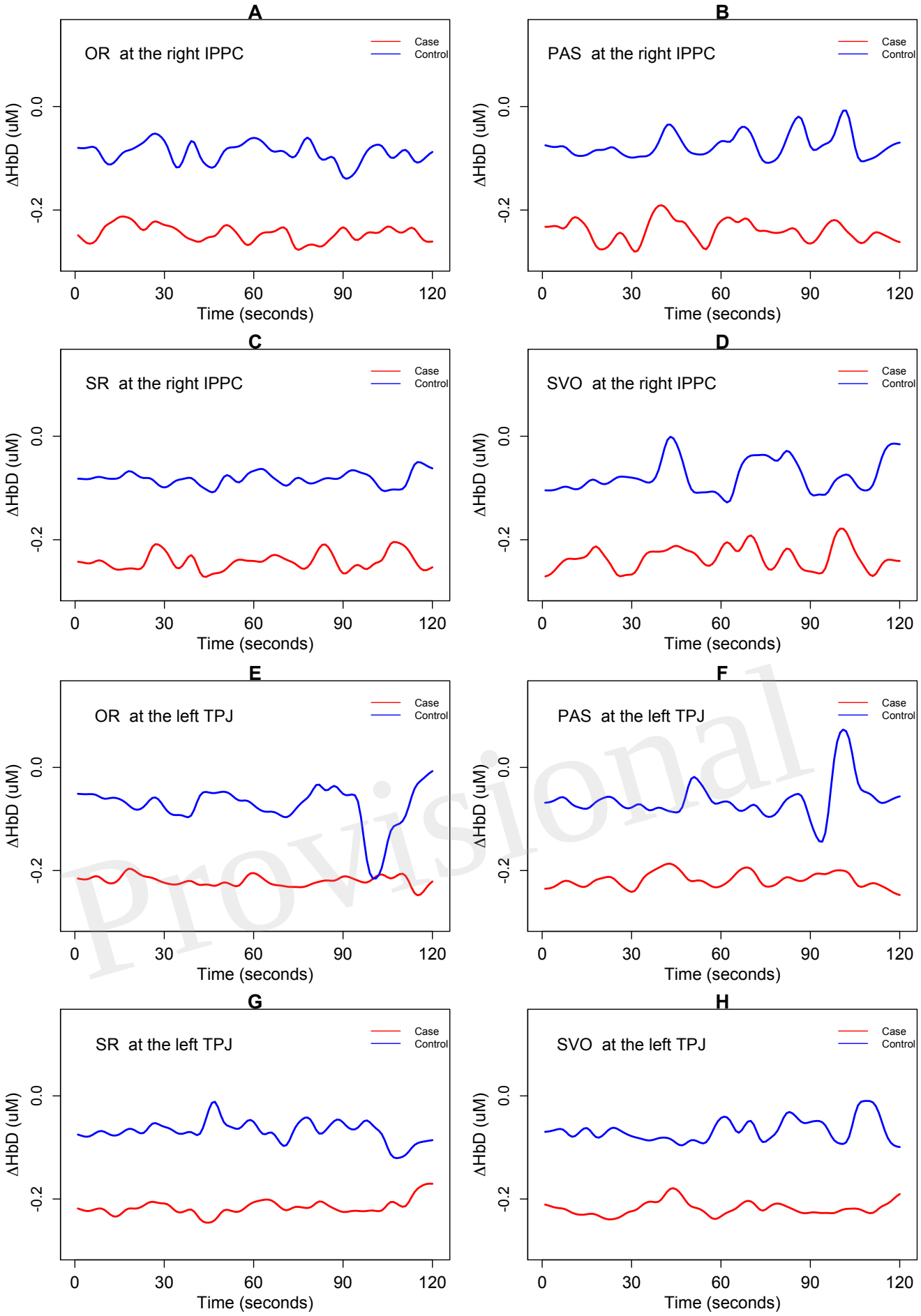

Figure 6: The mean trends for ΔHbD (i.e. H_{01}). IPPC = *inferior posterior parietal cortex*; TPJ= *temporal parietal junction*. OR = object relative clause sentences; SR = subject relative clause sentences; SVO = subject-verb-object sentences; and PAS = passive sentences. A: the mean trajectories of ΔHbD for OR syntax type at the right IPPC; B: the mean trajectories of ΔHbD for PAS syntax type at the right IPPC; C: the mean trajectories of ΔHbD for SR syntax type at the right IPPC; D: the mean trajectories of ΔHbD for SVO syntax type at the right IPPC; E: the mean trajectories of ΔHbD for OR syntax type at the left TPJ; F: the mean trajectories of ΔHbD for PAS syntax type at the left TPJ; G: the mean trajectories of ΔHbD for SR syntax type at the left TPJ; H: the mean trajectories of ΔHbD for SVO syntax type at the left TPJ.