Semantic and syntactic specialization during auditory sentence processing in 7-8-year-old children

Jin Wang[a,*], **Neelima Wagley**[a], **Mabel L. Rice**[b], **James R. Booth**[a]

[a]Department of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee, USA.

[b]Child Language Doctoral Program, University of Kansas, Lawrence, Kansas, USA.

Abstract

Previous studies indicate that adults show specialized syntactic and semantic processes in both the temporal and frontal lobes during language comprehension. Neuro-cognitive models of language development argue that this specialization appears earlier in the temporal than the frontal lobe. However, there is little evidence supporting this proposed progression. Our recently published study (Wang, Rice & Booth, 2020), using multivoxel pattern analyses, detected that children as young as 5 to 6 years old exhibit specialization and integration in the temporal lobe, but not the frontal lobe. In the current study, we used the same approach to examine semantic and syntactic specialization in children ages 7 to 8 years old. We found support for semantic specialization in the left middle temporal gyrus (MTG) for correct sentences and in the triangular part of the left inferior frontal gyrus (IFG) for incorrect sentences. We also found that the left superior temporal gyrus (STG) played an integration role and was sensitive to both semantic and syntactic processing during both correct and incorrect sentence processing. However, there was no support for syntactic specialization in 7- to 8-year-old children. As compared to our previous study on 5- to 6-year-old children, which only showed semantic specialization in the temporal lobe, the current study suggests a developmental progression to semantic specialization in the frontal lobe. This project represents an important step forward in testing neuro-cognitive models of language processing in children.

Keywords

Sentence Comprehension; Language Specialization; Multivoxel Pattern Analysis; Children

1. Introduction

Cortical regions begin at birth with broad functionality, but during development, some regions become more specialized to a narrow function. Neuro-developmental disorders are often characterized as having delayed processes of specialization or as having atypical patterns of specialization (see Interactive Specialization account in Johnson, 2011). Therefore, accurately characterizing brain specialization in developing children is important as it will not only promote our understanding of typical brain development but also provide a neural reference to determine what is different in children with neuro-developmental disorders.

A language comprehension model, proposed by Friederici (2012) and supported by previous studies (e.g., Newman et al., 2003; Hagoort & Indefrey, 2014), argues for specialized semantic and syntactic processing in the adult brain during language comprehension. According to this model, the opercular part (BA44) of the left inferior frontal gyrus (IFG) is specialized for syntactic processing (e.g., Hagoort & Indefrey, 2014; Friederici, 2018; Zaccarela, Schell, & Friederici 2017). The triangular part (BA45) of the left IFG and the left middle temporal gyrus (MTG) are specialized for semantic processing (e.g., Goucha, & Friederici, 2015; Hagoort & Indefrey, 2014; Binder et al., 2009). The superior temporal gyrus (STG) integrates the syntactic information from the opercular part of IFG and lexical semantic information from MTG (e.g., Zaccarella et al., 2017; Bornkessel et al., 2005). While analyzing the developing brain, Skeide and Friederici (2016) proposed that bottom-up processing for lexical-semantic and morphosyntactic categorization in the temporal lobe develops earlier than top-down processing for sentence-level semantics and syntax in the frontal lobe. The frontal lobe is thought to develop later and more gradually, continuing into young adulthood. However, it remains unclear when the different language regions specified in Friederici's (2012) model begin to show specialization and integration during language comprehension.

Only a few previous functional magnetic imaging (fMRI) studies have manipulated both semantic and syntactic factors to examine how the brain is specialized during auditory sentence processing in developing children. Most studies used univariate analysis and found that there were no functional dissociations in young children during complex sentence processing. Adult-like syntactic specialization during auditory sentence processing did not occur until children were 9-10 years old (e.g., Brauer & Friederici, 2007; Skeide, Brauer, & Friederici, 2014; Wu et al., 2016). Specifically, Brauer and Friederici (2007) found that 5-6-year-old children showed largely overlapping activation for syntactic violated, semantic violated and correct sentences, with the exception of greater activation in the opercular part of IFG (BA44) during the syntactic violated sentences. However, this greater activation may be because these sentences were more difficult. Wu and colleagues (2016) tested 5-6-year-old children and adults using a sentence-picture matching task, with a design of 2 syntactic complexity (subject-initial vs. object-initial sentences) by 3 animacy hierarchy (animate subject + inanimate object vs. animate subject + animate object vs. inanimate subject + animate object sentences). They found that 5-6-year-old children did not show adult-like functional dissociation. There was only a main effect of animacy in the triangular part of IFG, suggesting 5-6-year-old children relied on semantic cues when

processing auditory sentences. Skeide and colleagues (2014) tested children in several age groups (3-4, 6-7 and 9-10 years old) using a sentence-picture matching task, with a design of 2 syntactic complexity (sentences with subject-initial vs. object-initial clauses) by 2 plausibility (plausible vs. implausible sentences). They found that children aged 3-7 years old did not show adult-like specialization for semantic and syntactic processing, whereas 9-10-year-old children showed adult-like specialization for syntactic processing in the frontal lobe.

All of these previous studies on children used univariate analysis, but our recently published study (Wang, Rice & Booth, 2020) used multivoxel pattern analysis (MVPA) to examine the emergence of language specialization in 5-6-year-old children. One limitation of conventional univariate analysis is that it averages brain activation across voxels and therefore ignores the covariance of a voxel with other voxels. This covariance conveys information about cognitive representations/processes. MVPA provides a way to examine the information contained in a distributed pattern of activity (Mahmoudi et al., 2012) and has been shown to be more sensitive than univariate analysis in detecting subtle differences during language processing (e.g., Formisano et al., 2008; Mahon & Caramazza, 2010; Allen et al., 2012). In contrast to previous research using complex sentences, we designed relatively simple semantic and syntactic tasks to dissociate the two processes during auditory sentence processing. Adopting the MVPA approach from Haxby et al. (2001), we correlated the patterns of activation within versus across tasks. If the within-semantic task but not the within-syntactic task correlations are higher than the across-task correlations, this indicates that the brain region is only sensitive to semantic processing and distinguishes it from syntactic processing. The same logic applies to the analysis of syntactic processing if the within-syntactic task but not the within-semantic task correlations are higher than the across-task correlations. In addition, if the within-semantic task and within-syntactic task correlations are both higher than the across-task correlations, and there is no difference between the within-semantic task and within-syntactic task correlations, then this suggests an integration region sensitive to both semantic and syntactic processing. We found that the STG and MTG in 5-6-year-old children have developed distinct functions. The MTG was specialized for semantic processing, as evidenced by more similar activity patterns within the semantic task compared to across tasks, whereas there was no difference between the patterns within the syntactic task compared to across tasks. The STG, however, was sensitive to both semantic and syntactic information. This was demonstrated by more similar activity patterns for both within semantic and syntactic tasks compared to across tasks with no difference observed between the two within-task pattern similarities. These results suggest that semantic specialization and integration of semantics and syntax in the young children's brain (i.e., the temporal cortex) emerges as early as 5-6 years old, even when they have not fully mastered syntactic processes such as the finiteness of verbs (e.g., Rice and Wexler 2001). However, this study did not find semantic or syntactic specialization in the frontal lobe, which is consistent with the developmental model proposed by Skeide and Friederici (2016), in which they argue that language specialization occurs in the temporal lobe first and later in the frontal lobe. We also used univariate analysis in Wang et al. (2020), however, we did not find evidence for semantic or syntactic specialization in either the temporal or the frontal lobe. The lack of effects for univariate analyses, but the presence of effects for

MVPA suggest that semantic and syntactic specialization is best thought of as a distributed pattern across voxels.

In addition to the novelty of using MVPA (Wang et al., 2020), we only compared semantic and syntactic processing using ecologically valid correct sentences. A great deal of previous research on the development of language comprehension has focused on sentences with semantic and syntactic errors or non-canonical syntactic structure (e.g., Brauer & Friederici, 2007; Skeide et al., 2014). However, Davis and Rodd (2011) argued that anomalous sentences cause comprehension failure, the consequences of which are difficult to interpret because of the large variability. Schneider and Maguire (2019) have suggested the necessity of studying grammatically correct sentences to provide a more realistic view of how children process language in their everyday life. Previous studies have suggested that adult-like syntactic specialization in the frontal lobe for non-canonical grammatical sentences appears in 9- to 10-year-old children (e.g., Skeide & Friederici, 2016), but that canonical grammatical sentence processing appears to mature earlier. For example, Schneider et al. (2016) found that 10- to 12-year-old children differed from adults in comprehending grammatically incorrect sentences, but there were small differences for grammatically correct sentences. Because previous research used incorrect or non-canonical sentences (e.g., Brauer & Friederici, 2007; Skeide et al., 2014; Skeide & Friederici, 2016), and because our previous study with correct sentences in 5-6-year-old children (Wang et al., 2020) did not show syntactic specialization, it is crucial to use grammatically correct sentences in 7-8-year-old children to determine when syntactic specialization emerges.

Our previous study examined 5- to 6-year-old children, so it is important to extend the investigation of specialization to 7- to 8-year-old children. Children appear to shift from the optional use of morphosyntactic markers, producing sentences like "*Patsy walk(ed) home yesterday", to the obligatory use of adult grammar in this age range. As shown in the norms of the Test of Early Grammatical Impairment (TEGI, Rice & Wexler, 2001), 5- to 6-year-old children are still developing their morphosyntactic skills, with an average accuracy of 80-85% at 5;0-5;5 years old to a near ceiling average accuracy of 93-98% at 6;6-6;11 years old. Our previous study (Wang et al., 2020) examined 5- to 6-year-old children but did not find adult-like syntactic specialization in the frontal lobe probably due to their immature grammar. 7- to 8-year-old children, however, should have made the shift and thus set a benchmark for the obligatory use of adult grammar. Therefore, examining whether 7- to 8-year-old children exhibit adult-like frontal syntactic specialization, using a similar design as we did in 5- to 6-year-old children, would shed light on our understanding of this critical child to adult grammar shift during language development.

In the current study, we aimed to examine whether syntactic and semantic specialization occurred in children aged 7-8 years old using the same experimental design and a similar multi-voxel pattern analytical approach as in Wang et al. (2020) with 5- to 6-year-old children. We used a semantic task which involved plausibility judgment and a syntactic task which involved grammaticality judgment. We only compared grammatically correct sentences in the syntactic task with semantically plausible sentences in the semantic task so that the patterns of activation reflected comprehension processes of sensible sentences rather than anomalous ones (Davis & Rodd, 2011). According to Friederici's (2012)

language comprehension model, we hypothesized that if 7-8-year-old children have adult like specialization:

1.   the left MTG would be specialized for semantic processing. We expected that the within-semantic correlations would be positive and significantly higher than the across-task correlations, whereas the within-syntactic correlations would not be different from 0 and there would be no difference between the within-syntactic and across-task correlations.

2.   the left STG would be involved in the integration of syntactic and semantic processing. We expected that the within-semantic and within-syntactic correlations would be both positive and significantly higher than the across-task correlations, and there would be no difference between the within-semantic and within-syntactic correlations.

3.   the opercular part of the left IFG would be specialized for syntactic processing. We expected that the within-syntactic correlations would be positive and significantly higher than the across-task correlations, whereas the within-semantic correlations would not be different from 0 and there would be no difference between the witliin-semantic and across-task correlations.

4.   the triangular part of the left IFG would be specialized for semantic processing. We expected that the within-semantic correlations would be positive and significantly higher than the across-task correlations, whereas the within-syntactic correlations would not be different from 0 and there would be no difference between the within-syntactic and across-task correlations.

2.   Materials and Methods

2.1.   Statement

This is a registered report for a secondary analysis of existing data. The data was drawn from a large longitudinal project, which followed children at various stages from 5-6 years old to 7-8 years old to 9-10 years old. No one in the lab had examined the data from the 7-8-year-old children. The project coordinator and undergraduate research assistants collected the data and inputted them into the database. These personnel were not authors of this registered report. The first author of this report assisted with a few scanning sessions for a handful of participants. During the scanning sessions, she only had access to the raw brain images displayed on the screen in the imaging center and checked the accuracy of each run to make sure that children were paying attention to the tasks. These observations had no implications for our research question. Other authors had not observed this data set.

2.2.   Power analysis

We used G*Power to calculate the needed sample size for the MVPA analysis for each region of interest (ROI) separately. For MTG, we re-calculated the paired sample t test for SCon vs. Across in MTG from our previous paper (Wang et al., 2020) on 5-6-year-old children which used the same experimental paradigm and a similar analytical approach. This resulted in $t(29)=4.702$ and $p<0.001$ with an effect size of 0.873. With tails=1, α err

prob=0.02, power=0.90, the needed sample size calculated by G*Power is 17. For STG, we also re-calculated the paired sample t tests for both semantic and syntactic effects in STG from the Wang et al. (2020) paper on 5-6-year-old children, respectively. This resulted in t(29)=3.382, p=0.002, effect size=0.628 for Gram vs. Across, and t(29)=5.987, p<0.001, effect size=1.112 for SCon vs. Across. With tails=1, α err prob=0.02, power=0.90, the needed sample size calculated by G*Power is 31 for Gram vs. Across and 12 for SCon vs. Across in STG. Because 5-6-year-old children have not developed language specialization in the frontal lobe as shown in our previous study (Wang et al., 2020) and others (Brauer & Friederici, 2007; Skeide et al., 2014; Wu et al., 2016), for the opercular part of IFG, we calculated power based on the only previous study (Newman et al., 2003) that used the same ROI and similar contrasts in adults. They found a double dissociation in the frontal lobe with an effect in IFG.oper for syntactic violated (noun-verb) sentences versus fixation baseline [F(1,12)=4.10, p<0.07]. Based on this, we calculated the paired sample t(12)=2.02, and effect size =0.584. Then in G*Power, with tails=1, α err prob=0.02, power=0.90, the needed sample size is 35. For the triangular part of IFG, we also calculated power based on Newman et al. (2003) study, in which they found a double dissociation in the frontal lobe with an effect in IFG.tri for semantic violated (extraneous verb) sentences versus fixation baseline [F(1,12)=5.63, p<0.04]. Based on that we calculated the paired sample t(12)= 2.373, and effect size=0.685. In G*Power, with tails=1, α err prob=0.02, power=0.90, the needed sample size is 26. The details of parameters used in G*Power to calculate the needed sample size can be seen in Table 1. We chose the largest number (35 participants) among all these power analyses as the number of participants needed to have enough power to examine all our hypotheses.

2.3. Procedure and experimental tasks

Children were recruited from the Austin, Texas metropolitan area. Informed consent was obtained from the parents. The Institutional Review Board approved all of the following procedures.

2.3.1. Questionnaires and Standardized Testing—Participants were given developmental history questionnaires completed by their parents and a series of screening tests. The screening tests included a handedness interview in which the children were asked to write, erase, pick up, open, and throw something, as well as the Diagnostic Evaluation of Language Variation (DELV) Part 1 Language Variation Status (Seymour et al. 2003). Standardized testing was then administered to assess their IQ and language abilities. This included the non-verbal subtest of the Kaufman Brief Intelligence Test, Second Edition (KBIT-2, Kaufman and Kaufman 2004), and the core language measure of the Clinical Evaluation of Language Fundamentals (CELF-5, Wiig, et al. 2013).

2.3.2. In-scanner Sentence Tasks

__Stimuli:__ All sentence stimuli in the syntactic task and the semantic task had the following structure: An optional carrier phrase ("Last week"/ "Every day") + subject and verb phrase (e.g., "She baked") + number and object (e.g., "two cakes"). The sentences included one of the following four verb forms: 1) Third person present tense (-s); 2) Present progressive copula (be); 3) Auxiliary verb (do); and 4) Simple past tense (-ed). Each condition had

five sentence stimuli for each verb form (see below for a description of conditions). Stimuli were matched across all conditions in each task in terms of the written word frequency (Balota et al. 2007; Masterson et al. 2003), the number used (one /two /three /four /five /six), the subject used (he/she/they), the number of syllables (6-8), and the frequency of "not" usage in the sentences. The auditory sentences were recorded in a sound insulated booth by using Audacity software. All sentences were read by one female native English speaker who was asked to briefly pause between phrases. All sentences were then segmented into two (subject phrase + object phrase) or three (i.e., carrier phrase + subject phrase + object phrase) sections. Consistent pauses (approximately 500 ms) were added in between phrases using Praat software so that all sentences were similar in their pacing.

Syntactic Task: In each trial, children heard one auditory sentence, presented binaurally through earphones. There were three conditions of sentence stimuli: grammatically correct (Gram), finiteness violation (FVio), and plurality violation (PVio) (examples, see Table 2). A carefully matched frequency-modulated white noise burst served as the auditory perceptual control (PC) condition. The children were asked, "does the way she speaks sound right?" They were instructed to respond to all trials as quickly and accurately as possible, using their right index finger for a yes response in the Gram condition, and using their right middle finger for a no response in PVio and FVio conditions. Children were asked to press the yes button with their right index finger whenever they heard the PC condition. Throughout the trial, a blue circle remained on the screen during the auditory stimuli presentation and turned yellow 1000ms before the trial ended to remind the participants to respond. The duration of each sentence was 2700ms to 4500ms. The duration of the response interval was 2300ms. To optimize the extraction of the hemodynamic response, inter-trial intervals of 0, 575, or 1150ms were added randomly in equal proportions, resulting in a duration of 5000ms to 7950ms for each trial. The length of trials was equated across conditions. The four conditions were pseudo-randomized so that there were no more than 5 same responses in a row. There were 20 trials for each condition, totaling 80 trials evenly divided into two runs. Each run lasted around 4.5 minutes.

The three sentence conditions in the syntactic task were designed according to the following standards. The plurality violation condition was defined as the mismatch between the number and object by either adding an "s" or omitting an "s" in the object noun word. The finiteness violation condition was defined as the inconsistency between the subject and verb phrase by either adding an inflection or omitting an inflection/auxiliary verb. The grammatically correct condition was defined as sentences without grammatical errors.

Semantic Task: Similar to the syntactic task, in each trial, children heard one auditory sentence, presented binaurally through earphones. There were three conditions of the sentence stimuli: strongly congruent (SCon), weakly congruent (WCon) and incongruent (InCon) (examples, see Table 2). A carefully matched frequency-modulated white noise burst served as the auditory perceptual control (PC) condition. The children were asked, "does the way she speaks make sense?" They were instructed to respond to all trials as quickly and accurately as possible by using the right index finger for a yes response in SCon, WCon conditions, and using the right middle finger for a no response in the InCon

condition. Children were asked to press the yes button with their right index finger whenever they heard the PC condition. The presentation procedure was exactly the same as the syntactic task. There were 20 trials for each condition, totaling 80 trials evenly distributed in two runs. Each run lasted approximately 4.5 minutes.

The three sentence conditions in the semantic task were designed according to the following standards. The two congruent conditions were based on the association strength values between the verb and the object as defined in the University of South Florida Free Association Norms (Nelson et al. 1998). The strongly congruent condition had an association of 0.28-0.81 (M=0.41, SD=0.12) between the verb and the object in the sentence. The weakly congruent condition had an association of 0.02-0.19 (M=0.11, SD=0.05) between the verb and the object in the sentence. In the incongruent condition, the verb and the object in the sentence had no semantic association.

The Gram and the SCon trials were chosen as the best stimuli to examine semantic and syntactic specialization. One reason was that both conditions require the same response (pressing the yes button), excluding the possible confounding factor that different responses might induce distinct brain activation patterns. The other reason for choosing the Gram and the SCon trials was that both conditions are correct sentences, thereby avoiding potential confusion in processing anomalous sentences (Davis & Rodd, 2011). In this study, we chose the strongly congruent condition (SCon) in the semantic task as a condition of interest rather than the weakly congruent condition because the former was more natural and semantically predictable. Children's response may be more variable during weakly congruent sentence processing due to different language experience (e.g., Schneider & Maguire, 2019).

Prior to taking part in the fMRI scanning session, participants were required to complete a mock scan session. They performed the same task in the mock scanner, in order to ensure that they understood the task and were acclimated to the scanner environment. Different stimuli were used in the mock and real scanning session. The real scanning took place within a month of the practice session. The two tasks were counterbalanced.

2.3.3.   Data Acquisition—Participants lay in the scanner with a response button box placed in their right hand. The participants viewed a screen via a mirror attached to the inside of the head coil. The visual dot was projected onto a screen to keep participants focused on the task so that they would respond in time. Participants wore earphones to hear the auditory stimuli and two ear pads were used to attenuate the scanner noise. The two runs of a task were usually acquired within one session. If participants failed to finish certain runs, we invited them back a second time soon after the first session. Overall time-in-scanner for one session was less than one hour.

Images were acquired using 3 T Siemens Skyra MRI scanner with a 64-channel head coil. The blood oxygen level dependent (BOLD) signal was measured using a susceptibility weighted single-shot echo planar imaging (EPI) method. Functional images were acquired with multiband. The following scan parameters were used: TE=30ms, flip angle=80, matrix size=128×128, FOV=256mm, slice thickness=2mm without gap, number of slices=56, TR=1250ms, Multi-band accel.factor=4. A high resolution, T1 weighted 3D image was

acquired. The following scan parameters were used: TR=1900ms, TE=2.34ms, matrix size=256×256, field of view=256mm, slice thickness=1mm, number of slices=192.

2.4. Planned statistical analysis

2.4.1.   Participant inclusion criteria—In our study, 194 7-8-year-old participants completed both semantic and syntactic tasks, regardless of the data quality. Participants included in the data analyses met the following criteria: (1) Primarily right-handed, defined as performing at least 3 out of 5 items using their right hand during the handedness interview; (2) Mainstream English speakers, defined as scoring 9 or more (out of 15) for 7-year-old children and 11 or more (out of 15) for 8-year-old children for the mainstream English items on the DELV. (3) No diagnosis of Attention Deficit Hyperactivity Disorder (ADHD), neurological disease, psychiatric disorder, learning disability or language disorder, as well as normal hearing and normal or corrected-to-normal vision, as reported in the questionnaires. (4) Normal IQ and language skills, defined as the standardized score of higher than 80 for both the KBIT-2 non-verbal subtest and the CELF-5 core language score. (5) Acceptable accuracies for each run during the in-scanner tasks, defined as the accuracy of the perceptual control condition being greater than 60%. The accuracies of the easiest experimental condition in each task (i.e., SCon and PVio) being greater than 40%. No evidence of a response bias, as indicated by an accuracy difference lower than 40% between the InCon and SCon conditions or between the PVio and Gram conditions. (6) Acceptable head movement during the in-scanner tasks, defined as participants having no more than 10% or 6 consecutive outlier volumes in each run (see more details in the preprocessing section). Three participants were excluded due to left handedness. Twelve participants were excluded due to language variations. Twelve participants were excluded due to low non-verbal IQ or core language scores. Thirty-five participants were excluded due to movement in the scanner and 56 participants were excluded due to accuracy criteria. After all these above screening criteria, 76 participants (46 girls, mean age 7.35 ± 0.30, range 7.0-8.3 years old) were included in the final analysis. Descriptive statistics for the in-scanner task performance are presented in Table 3.

2.4.2.   Data analysis proposal

Preprocessing: The SPM12 toolbox (Statistical Parametric Mapping) (http://www.fil.ion.ucl.ac.uk/spm) was used to analyze the data. First, all functional images were realigned to their mean functional image across runs. Then, the anatomical image was segmented and warped to the pediatric tissue probability map template (Wilke, et al., 2017) to obtain the transformation field. An anatomical brain mask was created by combining three segmentation products (i.e., grey, white and cerebrospinal fluid), and then applied to its original anatomical image to produce a skull-stripped anatomical image. After that, we co-registered the mean functional image and all functional images to the skull-stripped anatomical image. All the functional images were then normalized to the pediatric template by applying the transformation field to them. We created the pediatric tissue probability map template by using CerebroMatic (Wilke, et al., 2017), a tool that makes SPM12 compatible pediatric templates with user-defined age, gender, and magnetic field. We chose the unified segmentation parameters estimated from 1919 participants described in the Wilke et al. (2017) (downloaded from https://www.medizin.uni-tuebingen.de/kinder/en/research/

neuroimaging/software/). We defined ages as 5.5-8 years old with one-month interval, two females and two males at each age interval with 3T scans, resulting in a sample size of 124 for our pediatric template. We used this template because it is the same as that we used for 5-6-year-old children (Wang et al., 2020), which allows for comparability. Art-Repair (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html) was used to identify outlier volumes, defined as those with volume-to-volume head movement exceeding 1.5 mm in any direction, head movement greater than 5 mm in any directions from the mean functional image or deviations of more than 4% from the mean global signal. To control for any movement effect on brain signal, the outlier volumes identified by Art-Repair were then repaired by interpolation between the nearest non-outlier volumes and de-weighted in the first level modeling (Mazaika et al. 2009). Participants having more than 10% or more than 6 consecutive outlier volumes in each run were excluded from the current study.

Regions of Interest: Based on Friederici's (2012) language comprehension model, we made 4 language regions of interest, namely, the opercular and the triangular part of the left IFG, the left STG, and the left MTG. The four language masks of interest were defined as the overlap between functional activation map at the group level (voxel wise threshold p=1) and anatomical mask of interest created by using the anatomical automatic labeling (AAL) atlas in the WFU PickAtlas tool (http://www.nitrc.org/projects/wfu_pickatlas). Because the AAL atlas is based on the adult brain, we warped the T1 structure of the AAL atlas to our pediatric T1 template using AFNI's 3dQwarp non-linear coregistration and then applied this transformation to the AAL atlas using AFNI's 3dNwarpApply. In this way, anatomical atlas masks were aligned with our pediatric T1 template.

Multi-voxel Pattern Analysis: Unsmoothed data was used to perform both feature selection and multi-voxel pattern analysis. For feature selection, we first estimated a traditional GLM with eight conditions from both the semantic and syntactic tasks. A high-pass filter with a cutoff of 128 sec and an SPM default artificial mask threshold of 0.5 was applied. Six movement parameters estimated from the realignment step were entered to control for movement effects. We included all trials, both correct and incorrect responses, in our model. In this way, there were an equal number of trials to compare across conditions and a large number of trials to increase power. The potential drawback of using all trials is that error responses may involve difference processes than correct responses. However, this potential confound is likely small because several previous studies have shown similar effects when modeling only correct trials versus modeling all trials (e.g., Hammer et al., 2015; Demir, Prado & Booth, 2015). In addition, our previous study (Wang et al., 2020) showed that the specialization effects stayed the same when children's behavioral performance (i.e., accuracy) was accounted for. After the GLM modeling, the contrast maps for all sentence conditions versus perceptual control conditions across tasks were generated. We chose the top 250 most activated voxels for the contrast within each of the four language ROIs (i.e., the opercular and the triangular part of the left IFG, the left MTG as well as the left STG) separately regardless of significance ($p = 1$). These top 250 voxels served as the features (voxels) that were the most sensitive to auditory sentence processing for the following multi-voxel pattern analysis. The overlap among participants' individualized top250-voxel

ROI within the masks of the left MTG, STG, opercular part of the left IFG, and triangular part of the left IFG are plotted in Figure 3 on the left.

Figure 1 illustrates the procedure of the multi-voxel pattern analysis (MVPA). Using a similar approach to Haxby et al. (2001), we compared the within-task and across-task correlations in the top 250 voxels within each ROI in order to examine semantic and syntactic specialization. For each task, we had two runs (namely, run1 and run2). We estimated the GLM for each run separately with four conditions (either Gram, FVio, PVio, PC or SCon, WCon, InCon, PC) as regressors of interest. Six movement parameters from the realignment step were entered in GLM as nuisance factors to control for movement effects. Then the contrast t-maps for Gram minus PC or SCon minus PC were generated for each run. The t-values in each voxel from the top 250 voxels within each ROI were then extracted using 3dMaskDump in AFNI toolbox (https://afni.nimh.nih.gov/). The within-semantic task correlation for each participant was calculated by correlating the t-values of the top 250 voxels for SCon minus PC in the semantic task run1 with the t-values of the top 250 voxels for SCon minus PC in the semantic task run2. In the same way, the within-syntactic task correlation for each participant was calculated by correlating the t-values of the top 250 voxels for Gram minus PC in the syntactic task run1 with the t-values of the top 250 voxels for Gram minus PC in the syntactic task run2. As for the across task correlation, the t-values of the top 250 voxels for Gram minus PC in either run1 or run2 of the syntactic task were correlated with the t-values of the top 250 voxels for SCon minus PC in either run1 or run2 of the semantic task, resulting in 4 between task correlations. The across-task correlation for each participant was then calculated by averaging the 4 between task correlations. We performed one sample t tests after that to examine the positivity of the pattern similarities within a task. We also conducted paired sample t tests to compare the within-semantic and within-syntactic correlations with across task correlations in each region to test our hypotheses. Details for the interpretations given different outcomes are illustrated in Table 1 and Figure 2. According to Friederici's (2012) language comprehension model, if 7-8-year-old children showed adult-like language specialization, we should observe (1) semantic specialization in MTG, (2) semantic and syntactic integration in STG, (3) syntactic specialization in the opercular part of IFG and (4) semantic specialization in the triangular part of IFG.

2.5 Exploratory analyses

2.5.1. Parallel MVPA analyses using incorrect sentences—Because children may process incorrect sentences differently from correct sentences (Schneider et al., 2016), to explore language specialization during incorrect sentence processing, parallel MVPA analyses were conducted using the finiteness violation condition (FVio) in the syntactic task and the incongruent condition (InCon) in the semantic task. We chose FVio condition rather than the plurality violation (PVio) condition in the syntactic task because it tapped into a core morphosyntactic skill (Rice and Wexler, 2001). We chose InCon condition in the semantic task as anomalous sentence manipulations have been extensively studied in the Event Related Potential (ERP) literature (e.g., N400 effect) as a reliable neural measure of semantic processing (e.g., Kutas & Hillyard, 1980; Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008).

2.5.2. Univariate analyses using both correct and incorrect sentences—

Because MVPA investigates the covariance among voxels, whereas univariate analysis examines the amount of activation at each voxel, we added univariate analysis to investigate task differences, which allows for a comparison to other studies. Data was smoothed with a 6-mm isotropic Gaussian kernel after normalization. Art-Repair was used to interpolate outlier volumes after smoothing. First-level analyses was performed on the smoothed data with a traditional GLM. The onsets of eight conditions from both the semantic and syntactic tasks were entered as regressors of interest and six movement parameters estimated from the realignment step were entered as regressors of no interest to control for movement effects. All trials were included, and outlier volumes identified by Art-Repair were de-weighted (Mazaika et al. 2009) in our model. A high-pass filter with a cutoff of 128 sec and an SPM default artificial mask threshold of 0.5 was applied. After the modeling, contrast maps for Gram minus PC, SCon minus PC, (Gram minus PC) > (SCon minus PC), and (SCon minus PC) > (Gram minus PC) were generated to observe brain activation for each task and syntactic or semantic specialized activation in each participant during correct sentence processing. Contrast maps for FVio minus PC, InCon minus PC, (FVio minus PC) > (InCon minus PC), (InCon minus PC) > (FVio minus PC) were generated to observe brain activation for each task and syntactic or semantic specialized activation in each participant during incorrect sentence processing. One-sample t test was then conducted in the second-level analysis for each contrast to generate group-level activation maps.

Statistical significance for the group level analysis within the combined four language masks (12,661 voxels) was defined using Monte Carlo simulations using AFNI's 3dClustSim program (see http://afni.nimh.nih.gov/). 3dClustSim carries out a 10,000 iteration Monte Carlo simulation of random noise activations at a particular voxel-wise alpha level within a masked brain volume. Following the suggestions made by Eklund, Nichols, and Knutsson (2016) regarding the inflated statistical significance achieved using some packages, we used 3dFWHMx to calculate the smoothness of the data for each participant, using a spatial autocorrelation function, and then averaged those smoothness values across all participants. This average smoothness value was then entered into 3dClustSim to calculate the cluster size needed for significance. The threshold for the size of a significant cluster within the language mask was 25 voxels at voxel-wise threshold at $p < .001$ uncorrected and cluster-wise threshold at $p < .05$ FWE corrected. In addition, brain activation maps at the whole brain level are displayed in Figure S1 and Table S1. The threshold for the size of a significant cluster at the whole brain level was 81 voxels at height threshold of $p < .001$ uncorrected and cluster-wise threshold of $p < .05$ FWE corrected.

2.6 Data and code availability

This study was pre-registered with the Open Science Framework. The original proposal can be found at https://osf.io/8jqc7/. The raw data of the entire project including experimental stimuli, standardized testing, in-scanner performance, as well as anatomical and functional imaging is shared on OpenNeuro.org (see data descriptor in Wang et al., 2021, https://openneuro.org/datasets/ds003604/versions/1.0.2). The laboratory log, the exact participant ID and run used, as well as the data analysis code

implemented in the current study are shared on GitHub at https://github.com/wangjinvandy/ Syntactic_Semantic_Specialization_7_8_yo.

3. Results

3.1 Results of planned MVPA using correct sentences

First, the assumptions test was conducted before examining the hypotheses of the current study. Using Shapiro-Wilk normality test, we found that the across-task correlations in the left MTG as well as in the triangular and opercular part of the left IFG did not follow a normal distribution ($ps < .05$). In addition, the within-syntactic correlations in the left STG did not follow a normal distribution ($p = .037$). Two outliers in the across-task correlations in the opercular part of the left IFG, one outlier in the across-task correlations in the triangular part of the left IFG, and one outlier in the across-task correlations in the left STG were deleted. After that, pair-wise comparisons using either paired sample t tests or nonparametric Wilcoxon signed-ranks tests, together with the positivity test of the within-task correlations, were performed to examine language specialization within each ROI.

In the left MTG (see Figure 3A in the middle), both the within-syntactic and the within-semantic correlations were significantly greater than the across-task correlations [$z = -5.323$, $p < .001$; $z = -5.893$, $p < .001$, Bonferroni corrected]. In addition, the within-semantic correlations were significantly higher than the within-syntactic correlations [$t(75) = -2.954$, $p = .033$, Bonferroni corrected]. One-sample t test showed that both the within-syntactic and the within-semantic correlations were significantly greater than 0 [$t(75) = 4.980$, $p < .001$; $t(75) = 6.833$, $p < .001$, Bonferroni corrected]. These results provide weak evidence for semantic specialization within the left MTG during correct sentence processing. In the left STG (see Figure 3B in the middle), both the within-syntactic and the within-semantic correlations were significantly greater than the across-task correlations [$z = -5.623$, $p < .001$; $t(75) = 8.532$, $p < .001$, Bonferroni corrected]. However, there was no significant difference between the within-syntactic and the within-semantic correlations [$z = -1.532$, $p = .375$, Bonferroni corrected] and both were significantly greater than 0 [$t(75) = 11.335$, $p < .001$; $t(75) = 14.585$, $p < .001$, Bonferroni corrected]. These results provide evidence for sensitivity to both semantic and syntactic processing in the left STG but no evidence of specialization during correct sentence processing.

In the opercular part of the left IFG (see Figure 3C in the middle), we found no differences among the within-syntactic correlations, the within-semantic correlations, and the across-task correlations [within-syntactic versus across-task: $z = -1.527$, $p = .381$; within-semantic versus across-task: $z = -1.710$, $p = .261$; within-syntactic versus within-semantic: $t(75) = -.193$, $p = 1.000$; Bonferroni corrected]. In addition, both the within-syntactic and the within-semantic correlations were significantly lower than 0 [$t(75) = -4.131$, $p < .001$; $t(75) = -4.148$, $p < .001$, Bonferroni corrected]. We found similar results in the triangular part of the left IFG (see Figure 3D in the middle). There were no significant differences among the within-syntactic correlations, the within-semantic correlations, and the across-task correlations [within-syntactic versus across-task: $z = -1.405$, $p = .480$; within-semantic versus across-task: $z = -2.107$, $p = .105$; within-syntactic versus within-semantic: $t(75) = -1.079$, $p = .852$; Bonferroni corrected]. In addition, both the within-syntactic and the

within-semantic correlations were significantly less than 0 [$t(75) = -5.551$, $p < .001$; $t(75)$ $= -3.635$, $p < .001$, Bonferroni corrected]. These results provide no evidence for sensitivity to either syntactic or semantic processing in both the opercular and the triangular part of the left IFG during correct sentence processing.

3.2   Results of exploratory analyses

3.2.1.   Parallel MVPA using incorrect sentences—Using Shapiro-Wilk normality test, we found that the across-task correlations in the triangular part of the left IFG did not follow a normal distribution ($p = .026$). One outlier in the within-syntactic correlations and two outliers in the across-task correlations in the opercular part of the left IFG were deleted. In addition, two outliers in the across-task correlations in the triangular part of the left IFG and one outlier in the left MTG were deleted. After that, pairwise comparisons using either paired sample t tests or nonparametric Wilcoxon signed-ranks tests, together with the positivity test of the within-task correlations, were performed to examine language specialization within each ROI.

In the left MTG (see Figure 3A on the right), both the within-syntactic and the within-semantic correlations were significantly greater than the across-task correlations [$t(74) = 4.343$, $p < .001$; $t(74) = 6.656$, $p < .001$, Bonferroni corrected]. However, we did not find significant difference between the within-syntactic and the within-semantic correlations [$t(75) = -1.332$, $p = .561$, Bonferroni corrected]. One-sample t test showed that both the within-syntactic and the within-semantic correlations were significantly greater than 0 [$t(74) = 4.222$, $p < .001$; $t(75) = 6.198$, $p < .001$, Bonferroni corrected]. Similarly, in the left STG (see Figure 3B on the right), both the within-syntactic and the within- semantic correlations were significantly greater than the across-task correlations [$t(75) = 7.880$, $p < .001$; $t(75)$ $= 7.814$, $p < .001$, Bonferroni corrected]. There was no significant difference between the within-syntactic and the within-semantic correlations [$t(75) = .436$, $p = 1.000$, Bonferroni corrected] and both were significantly greater than 0 [$t(75) = 12.533$, $p < .001$; $t(75) = 12.406$, $p < .001$, Bonferroni corrected]. These results provide evidence for sensitivity to both semantic and syntactic processing in MTG and STG but no evidence of specialization during incorrect sentence processing.

In the opercular part of the left IFG (see Figure 3C on the right), we found no differences among the within-semantic correlations, the within-syntactic correlations, and the across-task correlations [within-syntactic versus across-task: $t(73) = 2.374$, $p = .060$; within-semantic versus across-task: $t(73) = 2.152$, $p = .105$; within-syntactic versus within-semantic: $t(74) = .086$, $p = 1.000$; Bonferroni corrected]. In addition, both the within-syntactic and the within-semantic correlations were significantly less than 0 [$t(74) = -4.678$, $p < .001$; $t(75) = -4.341$, $p < .001$, Bonferroni corrected]. These results provide no evidence for sensitivity to either syntactic or semantic processing in the opercular part of the left IFG during incorrect sentence processing. In the triangular part of the left IFG, only the within-semantic correlations were significantly greater than the across-task correlations [$z = -2.642$, $p = .024$, Bonferroni corrected]. There were no significant differences between the within-syntactic correlations and the across-task correlations [$z = -1.705$, $p = .264$, Bonferroni corrected]. These results provide moderate evidence for semantic specialization

in the triangular part of the left IFG during incorrect sentence processing. However, both the within-syntactic and the within-semantic correlations were significantly less than 0 [$t(75) = -3.604$, $p < .001$; $t(75) = -3.717$, $p < .001$, Bonferroni corrected].

3.2.2. Univariate analyses using both correct and incorrect sentences—The univariate voxel-wise results for each task and task comparisons within the combined language mask during both correct and incorrect sentence processing are shown in Table 4 and Figure 4. Both tasks activated similar language regions including the left STG, the left MTG, and/or the triangular part of the left IFG for correct (see Figure 4A) and incorrect (see Figure 4B) sentences. In the direct task comparison of correct sentences, there were no brain regions activated more for the syntactic task than the semantic task. However, we found a small but significant cluster in the left MTG showing greater activation for the semantic task as compared to the syntactic task (see Figure 4C). In direct task comparison of incorrect sentences, all language regions including the left STG, the left MTG, and the triangular and opercular part of the left IFG, showed greater activation for the syntactic task as compared to the semantic task. However, there were no clusters showing greater activation for the semantic task than the syntactic task (see Figure 4D). Brain activation maps at the whole brain level can be found in Figure S1 and Table S1.

4. Discussion

The current study examined semantic and syntactic specialization in 7- to 8-year-old children during spoken language processing. In our planned analysis using MVPA, we observed evidence for semantic specialization in the left MTG and semantic and syntactic integration in the left STG during correct sentence processing. However, we did not find evidence for specialization or sensitivity to semantic or syntactic processing in the frontal lobe. These results are consistent with the findings from our previous study on 5- to 6-year-old children, in which we also observed semantic specialization and integration only in the temporal lobe (Wang et al., 2020). Thus, the current study suggests that even at ages 7 to 8 years old, semantic and syntactic processes are not specialized in the frontal lobe when assessed using semantically plausible and grammatically correct sentences.

In our planned MVPA analysis, we noticed that the r values for the within-semantic correlations in MTG were relatively small. The within-semantic correlations were likely underestimated in our study because they were calculated across runs. Various factors across runs can induce differences in brain activation. Despite this, we still observed that the within-semantic correlations were significantly higher than the across-task correlations, suggesting that MTG is specialized for semantic processing. The difference between the within-semantic and across-task correlations in the current study was similar to a previous study which also used MVPA to examine lexical representation in adult brains (Fedorenko, Nieto-Castañon, & Kanwisher, 2012). In addition, we observed that the across-task correlations in STG were higher than that in MTG, which may be indicative of an overlap of semantic and syntactic processing in STG. However, this across-task finding may be related to phonological processing. In the literature, STG has often been shown as a region sensitive to phonological processing of auditory words (e.g., Weiss, Cweigenberg, & Booth, 2018; Wang, Yamasaki, Weiss, & Booth, 2021). However, we found that within-

task correlations were higher than the across-task correlations and there was no difference between the two within-task correlations. Therefore, our results suggest that the STG is reliably engaged in both semantic and syntactic processing and this activation is not related to the phonological component of auditory sentence processing.

We conducted two exploratory analyses to further examine language specialization in this age group. First, we applied the same MVPA approach to assess specialization while processing anomalous sentences. Specifically, we used sentences with an incongruency from the semantic task and sentences with a finiteness violation from the syntactic task. This allowed to test whether patterns of language specialization were different when processing incorrect as compared to correct sentences. Results indicate that the left STG and MTG were sensitive to both semantic and syntactic processing but showed no evidence for specialization in either case. Additionally, we found no evidence for specialization of or sensitivity to semantic or syntactic processing in the opercular part of the left IFG. The lack of syntactic specialization during incorrect sentence processing is consistent with the results using correct sentences. This finding also aligns with prior developmental studies which suggest that adult-like syntactic specialization in children does not appear until 9- to 10-years-old (Skeide et al., 2014; Skeide & Friederici, 2016). However, different from the results using correct sentences, which showed semantic specialization in the left MTG, we found moderate evidence of semantic specialization in the triangular part of the left IFG for incorrect sentences. Semantic specialization in the left IFG for incorrect sentences may be related to the greater cognitive demands required when processing anomalous sentences. According to the Memory, Unification, and Control (MUC) model by Hagoort (2013), frontal regions are crucial for unification operations with the anterior-ventral IFG involved in semantic processing and the posterior-dorsal IFG involved in phonological and syntactic processing. When participants encounter a semantically anomalous word, they may engage the semantic unification operations more to assist with comprehension (Davis & Rodd, 2001). Thus, our first exploratory analysis shows support for semantic specialization in the frontal lobe in 7- to 8-year-old children when assessed using incorrect, semantically anomalous, sentences.

In both the planned and first exploratory analyses, we observed that the within-task correlations in the frontal lobe were all slightly negative. Based on Haxby et al. (2001), within category stimuli typically elicit more positive correlations whereas stimuli from different categories elicit lower, or more negative, correlations. Thus, the observation of positive within-task correlations in the temporal lobe are expected, whereas the negative within-task correlations in the frontal lobe are unexpected. Given that the absolute values of the within-task correlations in the frontal lobe were quite small (rs < 0.1) and the overall minimal engagement of the frontal lobe during both tasks, the unexpected negativity of the within-task correlations could be due to a regression to the mean effect across runs. Specifically, the activation of the voxels having the highest values in run1 may have decreased in run2, whereas the activation of the voxels having the lowest values in run1 may have increased in run2. This regression effect likely reduced the within-task similarity and resulted in the negative correlations. To test this hypothesis, we compared run1 and run2's standard deviations of the top 250 active voxels in the frontal lobe. We found that the average standard deviations in run1 were always larger than run2 regardless of task or

condition, supporting the regression to the mean effect across runs. More generally, these results further suggest that the frontal lobe is still developing in children ages 7- to 8-years-old as indexed by the failure to robustly elicit similar patterns of brain activation across repetitions of the same language processing task. In addition, we noticed that behaviorally children had more difficulty performing the syntactic task as evidenced by significantly lower accuracy for both correct [$t(75) = -3.67. p < 0.001$] and incorrect [$t(75) = -7.75. p < 0.001$] sentence processing in the syntactic task than the semantic task. To examine if task difficulty affected the within-task and across-task correlations in our MVPA results, we correlated children's task accuracy with the within- and across-task correlations in each ROI for both correct and incorrect sentences. None of the correlational tests were significant after correcting for multiple comparisons in each ROI ($p = 0.05/3 = 0.017$). Therefore, our MVPA results, which were based on the within-task and across-task correlations, were likely not influenced by task difficulty.

In the second exploratory analysis, we used univariate analyses to evaluate language specialization during both correct and incorrect sentence processing. For correct sentences, no regions of interest (i.e., the left STG, MTG, or the triangular and opercular part of the left IFG) showed greater activation for the syntactic task than the semantic task. However, the left MTG showed greater activation for the semantic task than the syntactic task. Although not a double dissociation, this result is consistent with our MVPA findings that showed semantic specialization in the left MTG. Moreover, the lack of difference in the STG for the univariate analyses suggests comparable sensitivity to both semantic and syntactic processing. Overall, these findings are consistent with Friederici's (2012) language comprehension model, which suggests that the left MTG specializes for semantics whereas the left STG is an integrative region sensitive to both semantics and syntax.

For incorrect sentences, all language regions showed greater activation for the syntactic than the semantic task. However, no regions showed greater activation for the semantic task than the syntactic task. The greater activation for the finiteness violated sentences during the syntactic task likely reflects greater effort in recognizing the morphosyntactic errors as compared to the semantic plausibility errors, which parallels the children's behavioral task performance. Children had significantly lower accuracy for the FVio (mean = 72.4%) condition than the InCon (mean = 90.8%) condition [$t(75) = 7.754, p < .001$]. While the norms of Test of Early Grammatical Impairment (TEGI, Rice and Wexler, 2001) suggests that children ages 7 to 8 years old should have made the shift from the optional use of morphosyntactic markers to the obligatory use of adult grammar, it appears that recognizing morphosyntax is still computationally demanding. Indeed, Wagley et al. (2019) found that even proficient, native-English speaking adults continued to show effortful processing during finiteness errors evidenced by lower accuracy and slower response time as well as greater engagement of the dorsal IFG and STG associated with syntactic processing. Thus, in the current study, the greater activation in the opercular part of left IFG and the left STG for the finiteness violated sentences in 7-8-year-old children is likely a result of greater computational demand for morphosyntactic processing. However, different from the study by Wagley et al. (2019), we also observed greater activation in regions associated with semantic processing, such as the triangular part of the left IFG and the left MTG. Literature suggests that semantic processes emerge earlier than syntactic processes in development (see

review by Morgan et al., 2020) and that young children tend to utilize semantic cues to help with comprehending syntactically complex sentences (Wu et al., 2016; Schneider et al., 2019). Therefore, the 7- to 8-year-old children in our study likely relied on semantic processing in the brain to compensate for their immature syntactic skills during auditory sentence comprehension.

When comparing the current results of 7- to 8-year-old children with those from our previous study on 5- to 6-year-old children (Wang et al., 2020), a gradual maturation of the frontal lobe is implied. In the univariate analysis, 5- to 6-year-old children only showed activation in the temporal lobe but 7- to 8-year-old children additionally engaged the frontal lobe. The additional engagement of the frontal lobe suggests that older children effectively engage higher-level cognitive processes during sentence comprehension. For MVPA, 5- to 6-year-old children only showed semantic specialization in the temporal lobe for correct sentences. The 7- to 8-year-old children showed the same pattern and, in addition, exhibited semantic specialization in the triangular part of the IFG for incorrect sentences. Together, findings from our previous (Wang et al., 2020) and current study suggest a developmental progression that gradually engages frontal in addition to the temporal lobe during sentence comprehension. This is consistent with the neurocognitive model of language development proposed by Skeide and Friederici (2016), which argues that the temporal lobe develops earlier than the frontal lobe. However, one limitation of the current study is its cross-sectional design. Confounding factors such as sample size and cohort effects may influence brain activation across ages, which weaken the evidence for the developmental progression of semantic and syntactic specialization in children. Therefore, future studies using a longitudinal design are needed to address this issue.

Although our study suggests a developmental progression of increased frontal engagement, we did not find evidence for syntactic specialization in the frontal lobe in 7- to 8-year-old children using either univariate or multivoxel pattern analyses. Based on Friederici's (2012) language comprehension model, the opercular part of the left IFG should be the core region for syntactic processing in adults. Consistent with Friederici's (2012) model, Hagoort and Indefrey (2014) reviewed studies that directly compared sentences with high syntactic and high semantic demands. They showed that the left posterior IFG was reliably more active for syntactic than semantic processing, whereas the left anterior IFG was more active for semantic than syntactic processing in adults. However, the number of studies reviewed in Hagoort and Indefrey (2014) was small with only 6-10 studies directly comparing semantics and syntax. Rodd et al. (2015) conducted a meta-analysis with 54 fMRI studies which manipulated either semantics or syntax. Contrary to Friederici's (2012) language comprehension model, they did not find that the dorsal IFG showed greater activation for syntax than semantics. Rodd et al. (2015) pointed out that because the task manipulations in the semantic studies were different from those in the syntax studies, the interpretation for the semantic versus syntax comparisons is difficult. Goucha and Friederici (2015) progressively removed semantic information from a sentence and found that brain activation in the anterior left IFG (BA45) gradually diminished, but the dorsal left IFG (BA44) remained suggesting a role for syntactic processing. In addition, Zaccarella, Schell, and Friederici (2017) reviewed 19 studies which compared sentences with word lists to examine the neural basis of syntactic processing. They found that the dorsal left IFG (BA44) was only activated when the word

lists contained either functional or content words. They argued that when the word lists contain both functional and content words, syntactic merging may occur, resulting in a lack of activation in the dorsal left IFG (BA44). Some recent reviews and meta-analyses (e.g., Friederici, 2018; Walenski et al., 2019; Heard & Lee, 2020), which included various syntactic studies, suggest that syntactic processing does consistently engage the left dorsal IFG. However, because these studies did not directly compare syntactic with semantic processing, whether the dorsal IFG is the core location of syntactic specialization remains unclear.

Contrary to the assertation that a specific subpart of the frontal cortex is specialized for syntax (Friederici, 2018), others argue that syntactic processing is inseparable from processing meaning. For example, Fedorenko and her colleagues (2020) did not find stronger responses to syntactic processing than lexico-semantic processing, although some regions showed the opposite pattern. These results align with the lexicalist view of syntax, which argues that much of the structure of sentences is represented along with words in the lexicon, and thus syntax is not separable from lexical semantic processing. In line with this view, based on many studies on lesion-deficit mapping, Matchin and Hickok (2020) proposed that the left posterior MTG, a region often associated with semantic processing (Binder et al., 2009), is the core location for lexical-syntactic processing. However, Matchin and Hickok did not directly contrast syntactic with semantic processing, so it is unclear if the neural bases of syntactic processing can be distinguished from semantic processing during sentence comprehension. In summary, although semantic and syntactic specialization has been extensively studied in adults, it is still debated whether syntactic processing is separable from semantic processing and where it localizes in the brain.

The lack of consistency in the neural basis of syntactic processing in the previous literature could be related to the fact that syntactic effects are difficult to distinguish from semantic effects in natural language, given that syntactic changes usually alter the meaning of the expression (Pylkkänen, 2019). The sensitivity of designs used in previous studies may play an important role in determining whether a separation of syntactic from semantic processing can be found. With careful experimental designs using temporally precise measures, magnetoencephalography (MEG) studies have revealed double dissociations for semantic and syntactic processing. The posterior left temporal lobe and/or the left IFG was sensitive to syntactic processing, whereas the anterior left temporal lobe was sensitive to semantic processing (e.g., Matar et al., 2021; Law & Pylkkänen, 2021). However, due to the relatively poor spatial resolution in MEG, it is unclear which sub-regions in the left posterior temporal lobe and/or the left IFG are implicated in syntactic processing. Another reason for the inconsistency in the literature may be that multiple linguistic sub-operations involved during syntactic processing (e.g., merging, movement, re-analysis, etc.). Although the dorsal IFG appears to be commonly activated during syntactic processing, there were large variations in activation elicited by different sub-operations of syntactic processing (Heard & Lee, 2020). Further, task requirements may play a role in determining whether syntactic specialization can be detected. In the studies by Newman et al. (2003) and Friederici et al. (2003), where syntactic specialization was observed in the opercular part of the left IFG in adults, participants were asked to judge the correctness of sentences which forced them to pay attention to both sentence structure as well as meaning. However, in the study by

Fedorenko et al., (2020), which did not find syntactic specialization, the task requirements (i.e., passive reading, probe word detection, same or different meaning) all biased towards processing the meaning of the sentences. Unless the syntactic judgment is more demanding or explicit, syntactic processing in the brain may not be easily detected during natural language comprehension. In our study, we used two explicit tasks to maximally engage semantic and syntactic processing, but we still did not find syntactic specialization in 7- to 8-year-old children. Given that children this age continue to develop their syntactic skill, future studies in older children are needed to examine if syntactic specialization occurs later as proposed by developmental models of language comprehension.

In summary, the current study examined language specialization during auditory sentence processing in children 7- to 8-years-old. We found support for semantic specialization in the left MTG for correct sentences and in the triangular part of the left IFG for incorrect sentences. We also found that the left STG played an integration role and was sensitive to both semantics and syntax during both correct and incorrect sentence processing. We did not observe syntactic specialization in 7- to 8-year-old children. As compared to our previous study on 5- to 6-year-old children, which showed semantic specialization only in the temporal lobe, the current study suggests a developmental progression to semantic specialization in the frontal lobe, consistent with neuro-cognitive models for language development. Future studies in older children are needed to determine whether syntactic specialization can be observed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

6. References

Allen K, Pereira F, Botvinick M, & Goldberg AE (2012). Distinguishing grammatical constructions with fMRI pattern analysis. Brain and Language, 123(3), 174–182. [PubMed: 23010489]

Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, … Treiman R (2007). The English lexicon project. Behavior research methods, 39(3), 445–459. [PubMed: 17958156]

Binder JR, Desai RH, Graves WW, & Conant LL (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral Cortex, 19(12), 2767–2796. [PubMed: 19329570]

Bornkessel I, Zysset S, Friederici AD, Von Cramon DY, & Schlesewsky M (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. Neuroimage, 26(1), 221–233. [PubMed: 15862222]

Brauer J & Friederici AD (2007). Functional neural networks of semantic and syntactic processes in the developing brain. Journal of Cognitive Neuroscience, 19(10), 1609–1623. [PubMed: 18271736]

Davis MH, & Rodd JM (2011). Brain structures underlying lexical processing of speech: evidence from brain imaging. Lexical representation: a multidisciplinary approach. Berlin (Germany): Mouton de Gruyter, 197–230.

Demir ÖE, Prado J, & Booth JR (2015). Parental socioeconomic status and the neural basis of arithmetic: differential relations to verbal and visuo - spatial representations. Developmental Science, 18(5), 799–814. [PubMed: 25664675]

Eklund A, Nichols TE & Knutsson H (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences, 113(28), 7900–7905.

Fedorenko E, Blank IA, Siegelman M, & Mineroff Z (2020). Lack of selectivity for syntax relative to word meanings throughout the language network. Cognition, 203, 104348. [PubMed: 32569894]

Fedorenko E, Nieto-Castanon A, & Kanwisher N (2012). Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses. Neuropsychologia, 50(4), 499–513. [PubMed: 21945850]

Friederici AD (2012). The cortical language circuit: from auditory perception to sentence comprehension. Trends in Cognitive Sciences, 16(5), 262–268. [PubMed: 22516238]

Friederici AD (2018). The neural basis for human syntax: Broca's area and beyond. Current Opinion in Behavioral Sciences, 21, 88–92.

Friederici AD, Rüschemeyer SA, Hahne A, & Fiebach CJ (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cerebral Cortex, 13(2), 170–177. [PubMed: 12507948]

Formisano E, De Martino F, Bonte M, & Goebel R (2008). " Who" is saying" what"? Brain-based decoding of human voice and speech. Science, 322(5903), 970–973. [PubMed: 18988858]

Goucha T, & Friederici AD (2015). The language skeleton after dissecting meaning: a functional segregation within Broca's Area. Neuroimage, 114(1), 294–302. [PubMed: 25871627]

Hagoort P, & Indefrey P (2014). The neurobiology of language beyond single words. Annual Review of Neuroscience, 37, 347–362.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, & Pietrini P (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, 293(5539), 2425–2430. [PubMed: 11577229]

Heard M, & Lee YS (2020). Shared neural resources of rhythm and syntax: An ALE meta-analysis. Neuropsychologia, 137, 107284. [PubMed: 31783081]

Hammer R, Tennekoon M, Cooke GE, Gayda J, Stein MA, & Booth JR (2015). Feedback associated with expectation for larger-reward improves visuospatial working memory performances in children with ADHD. Developmental Cognitive Neuroscience, 14, 38–49. [PubMed: 26142072]

Johnson MH (2011). Interactive specialization: a domain-general framework for human functional brain development? Developmental Cognitive Neuroscience, 1(1), 7–21. [PubMed: 22436416]

Kaufman AS, & Kaufman NL (2004). Kaufman Brief Intelligence Test, Second Edition. Bloomington, MN: Pearson.

Kutas M, & Hillyard SA (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207(4427), 203–205. [PubMed: 7350657]

Kutas M, & Federmeier KD (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.

Law R, & Pylkkänen L (2021). Lists with and without syntax: A new approach to measuring the neural processing of syntax. Journal of Neuroscience, 41(10), 2186–2196. [PubMed: 33500276]

Lau EF, Phillips C, & Poeppel D (2008). A cortical network for semantics:(de) constructing the N400. Nature Reviews Neuroscience, 9(12), 920–933. [PubMed: 19020511]

Mahmoudi A, Takerkart S, Regragui F, Boussaoud D, & Brovelli A (2012). Multivoxel pattern analysis for FMRI data: a review. Computational and Mathematical Methods in Medicine, 2012(2012), 961257. [PubMed: 23401720]

Mahon BZ, & Caramazza A (2010). Judging semantic similarity: An event-related fMRI study with auditory word stimuli. Neuroscience, 169(1), 279–286. [PubMed: 20412836]

Masterson J, Stuart M, Dixon M, & Lovejoy S (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. British Journal of Psychology, 101(2), 221–242. [PubMed: 20021708]

Matchin W, & Hickok G (2020). The cortical organization of syntax. Cerebral Cortex, 30(3), 1481–1498. [PubMed: 31670779]

Matar S, Dirani J, Marantz A, & Pylkkänen L (2021). Left posterior temporal cortex is sensitive to syntax within conceptually matched Arabic expressions. Scientific Reports, 11(1), 1–14. [PubMed: 33414495]

Mazaika P, Hoeft F, Glover GH, Reiss AL (2009). Methods and software for fMRI analysis for clinical subjects. Human Brain Mapping Conference.

Morgan EU, van der Meer A, Vulchanova M, Blasi DE, & Baggio G (2020). Meaning before grammar: A review of ERP experiments on the neurodevelopmental origins of semantic processing. Psychonomic bulletin & review, 27(3), 441–464. [PubMed: 31950458]

Nelson DL, McEvoy CL, Schreiber TA (1998). The University of South Florida word association, rhyme, and word fragment norms [Database]. Retrieved from w3.usf.edu/FreeAssociation.

Newman SD, Just MA, Keller TA, Roth J, & Carpenter PA (2003). Differential effects of syntactic and semantic processing on the subregions of Broca's area. Cognitive Brain Research, 16(2), 297–307. [PubMed: 12668239]

Pylkkänen L (2019). The neural basis of combinatory syntax and semantics. Science, 366(6461), 62–66. [PubMed: 31604303]

Rice ML, Wexler K (2001). Rice/Wexler Test of Early Grammatical Impairment. San Antonio, TX: The Psychological Corporation.

Rodd JM, Vitello S, Woollams AM, & Adank P (2015). Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. Brain and Language, 141, 89–102. [PubMed: 25576690]

Schneider JM, Abel AD, Ogiela DA, Middleton AE, & Maguire MJ (2016). Developmental differences in beta and theta power during sentence processing. Developmental Cognitive Neuroscience, 19, 19–30. [PubMed: 26774879]

Schneider JM, & Maguire MJ (2019). Developmental differences in the neural correlates supporting semantics and syntax during sentence processing. Developmental Science, 22(4), e12782. [PubMed: 30525288]

Seymour HN, Roeper T, & De Villiers JG (2003). DELV: Diagnostic Evaluation of Language Variation: Screening Test. PsychCorp.

Skeide MA, & Friederici AD (2016). The ontogeny of the cortical language network. Nature Reviews Neuroscience, 17(5), 323–332. [PubMed: 27040907]

Skeide MA, Brauer J, & Friederici AD (2014). Syntax gradually segregates from semantics in the developing brain. Neuroimage, 100(15), 106–111. [PubMed: 24927987]

Wagley N, Perrachione TK, Ostrovskaya I, Ghosh SS, Saxler PK, Lymberis J, Wexler K, Gabrieli JDE, & Kovelman I (2019). Persistent Neurobehavioral Markers of Developmental Morphosyntax Errors in Adults. Journal of Speech, Language, and Hearing Research, 62(12), 4497–4508.

Wang J, Rice ML, & Booth JR (2020). Syntactic and Semantic Specialization and Integration in 5-to 6-Year-Old Children during Auditory Sentence Processing. Journal of Cognitive Neuroscience, 32(1), 36–49. [PubMed: 31596168]

Wang J, Lytle MN, Weiss Y, Yamasaki BL, & Booth JR (2021). A longitudinal neuroimaging dataset on language processing in children ages 5, 7, and 9 years old. Preprint at 10.31234/osf.io/tpndf

Weiss Y, Cweigenberg HG, & Booth JR (2018). Neural specialization of phonological and semantic processing in young children. Human Brain Mapping, 39(11), 4334–4348. [PubMed: 29956400]

Wang J, Yamasaki BL, Weiss Y, & Booth JR (2021). Both frontal and temporal cortex exhibit phonological and semantic specialization during spoken language processing in 7 - to 8 - year - old children. Human Brain Mapping, 42(11), 3534–3546. [PubMed: 33951259]

Walenski M, Europa E, Caplan D, & Thompson CK (2019). Neural networks for sentence comprehension and production: An ALE - based meta - analysis of neuroimaging studies. Human Brain Mapping, 40(8), 2275–2304. [PubMed: 30689268]

Wiig EH, Semel E, & Secord WA (2013). Clinical Evaluation Language Fundamentals (CELF-5) (5th ed.). San Antonio: Pearson.

Wilke M, Altaye M, Holland SK, & CMIND Authorship Consortium. (2017). CerebroMatic: a versatile toolbox for spline-based MRI template creation. Frontiers in Computational Neuroscience, 11, 5. [PubMed: 28275348]

Wu CY, Vissiennon K, Friederici AD, & Brauer J (2016). Preschoolers' brains rely on semantic cues prior to the mastery of syntax during sentence comprehension. Neuroimage, 126(1), 256–266. [PubMed: 26497266]

Zaccarella E, Schell M, & Friederici AD (2017). Reviewing the functional basis of the syntactic Merge mechanism for language: A coordinate-based activation likelihood estimation meta-analysis. Neuroscience & Biobehavioral Reviews, 80, 646–656. [PubMed: 28743620]
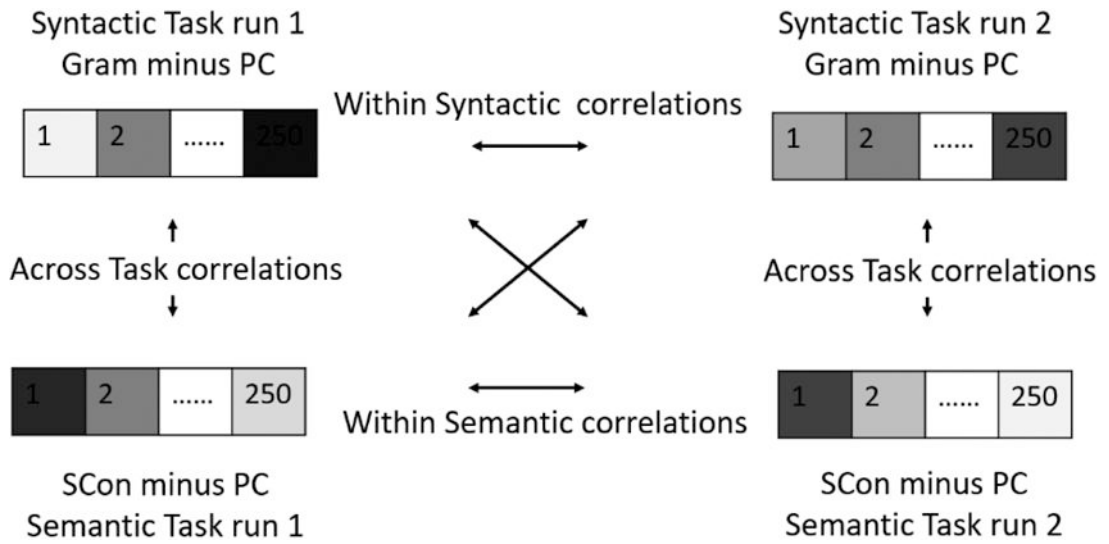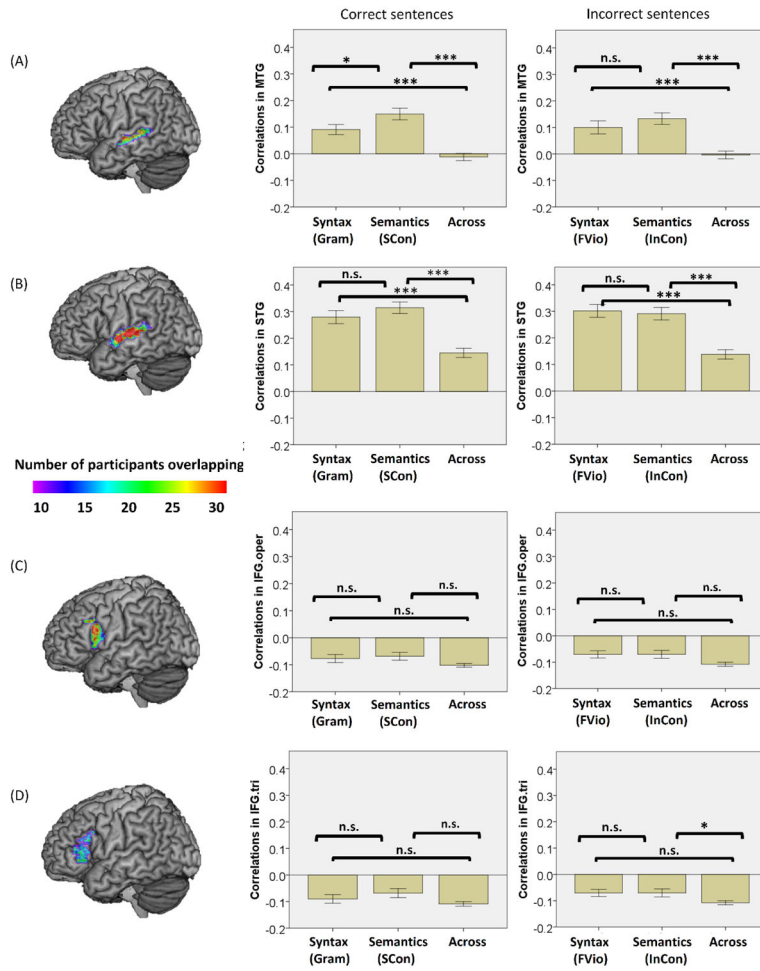
Figure 1.
The multi-voxel pattern analysis in this study based on Haxby, et al. (2001). Gram: grammatically correct; SCon: strongly congruent; PC: perceptual control. This figure was adopted from Wang et al. (2020).
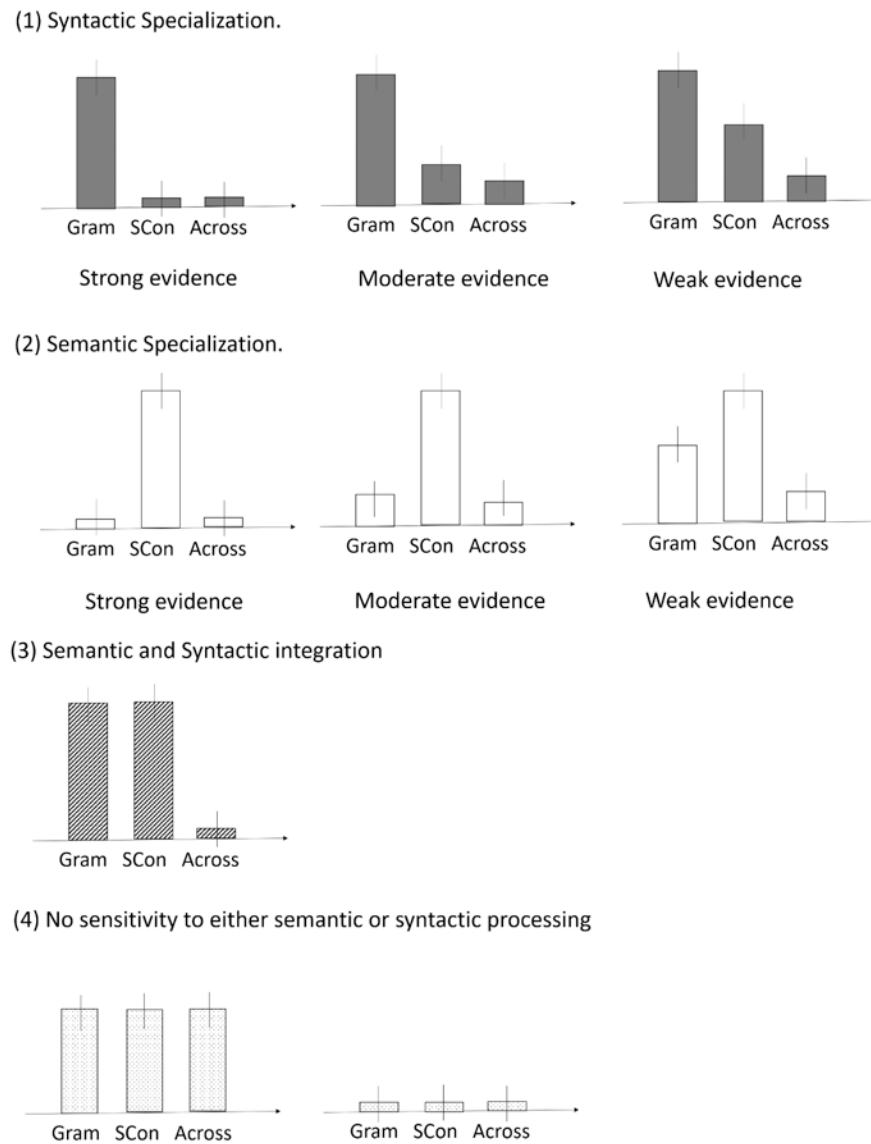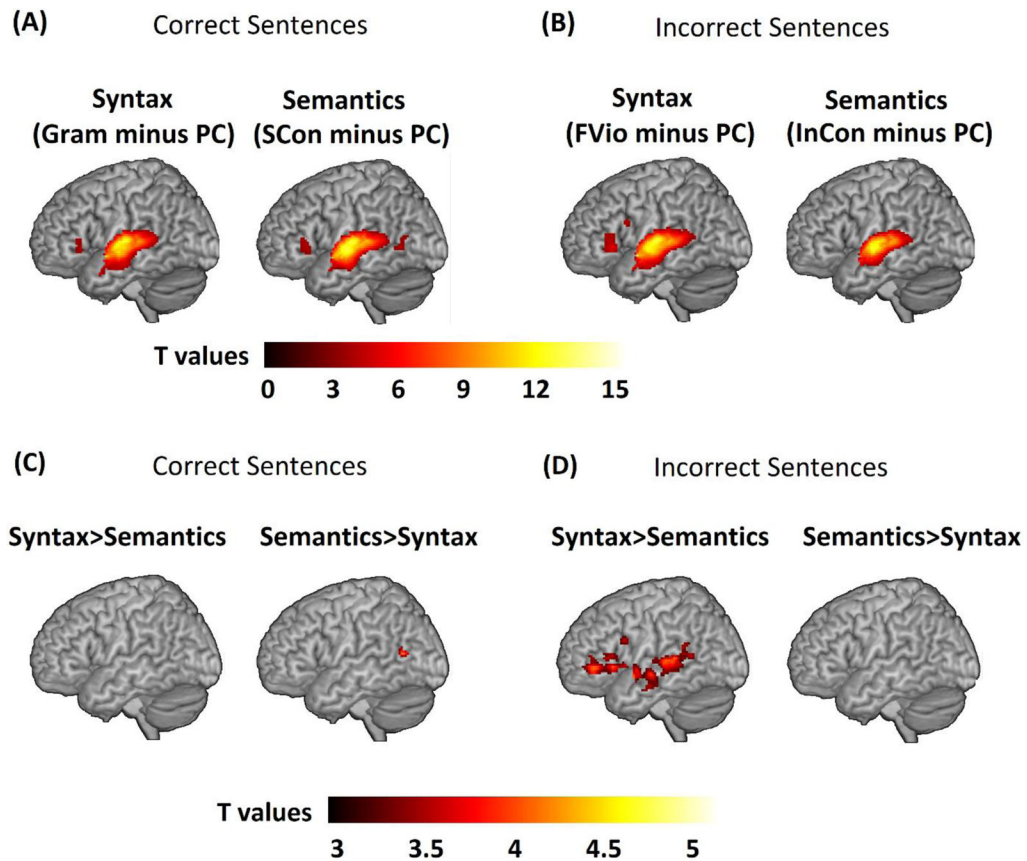
Figure 2.
Examples of different outcomes. Gram: within-syntactic correlations; SCon: within-semantic correlations; Across: across-task correlations.

Figure 3.
Statistics for the within-syntactic, the within-semantic, and the across-task correlations in the four language ROIs. (A) the left MTG. (B) the left STG. (C) the opercular part of the left IFG (IFG.oper). (D) the triangular part of the left IFG (IFG.tri). n.s. = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Bonferroni corrected. Gram and FVio refer to the within-syntactic correlations during correct and incorrect sentences, respectively; SCon and InCon refer to the within-semantic correlations during correct and incorrect sentences, respectively; Across refers to the across-task correlations.

Figure 4.

The univariate voxel-wise results for each task and task comparisons within the combined language mask. (A) Brain activation for each task for correct sentences. Gram minus PC refers to brain activation greater for grammatically correct (Gram) than perceptual control (PC) conditions in the Syntactic Task; SCon minus PC refers to brain activation greater for strongly congruent (SCon) than perceptual control (PC) conditions in the Semantic Task. (B) brain activation for each task for incorrect sentences. FVio minus PC refers to brain activation greater for finiteness violated (FVio) than perceptual control (PC) conditions in the Syntactic Task; InCon minus PC refers to brain activation greater for incongruent (InCon) than perceptual control (PC) conditions in the Semantic Task. (C) Task comparison results for correct sentences. (Gram minus PC) > (SCon minus PC) was used to evaluate brain activation for Syntax > Semantics. (SCon minus PC) > (Gram minus PC) was used to evaluate brain activation for Semantics > Syntax. (D) Task comparison results for incorrect sentences. (FVio minus PC) > (InCon minus PC) was used to evaluate brain activation for Syntax > Semantics. (InCon minus PC) > (FVio minus PC) was used to evaluate brain activation for Semantics > Syntax. Group maps thresholded at voxel-wise $p < 0.001$ uncorrected (T values > 3.2) and cluster-wise $p < 0.05$ corrected within the combined language mask (the left STG, MTG, and the triangular and opercular part of the left IFG). Clusters with size greater than 25 voxels are shown.

Table 1.

The proposed hypotheses, the power analyses, the proposed statistical tests and the interpretation given different outcomes.

Question	Hypothesis	Sampling plan (e.g., power analysis)	Analysis Plan	Interpretations given different outcomes
1. Is the middle temporal gyrus (MTG) specialized for semantic processing?	The within-semantic correlations > 0 and the within-semantic correlations > across-task correlations. In addition, the within-syntactic correlations = 0, and the within-syntactic correlations = across-task correlations.	We re-calculated a paired sample t test for SCon vs. Across in MTG from our previous paper (Wang et al., 2020) on 5-6-year-old children which used the same experimental paradigm and same analytical approach. This resulted in t(29)=4.702, p<0.001 effect size=t/sqrt(n)=0.873. In G*Power, we chose the dependent sample t test. With Tails=1, Effect size dz=0.873, α err prob=0.02, Power=0.90. The needed sample size is 17.	The analyses were the same for all four regions. All of the following steps were performed for each region. Assumptions Tests: (1) Normal distribution: Using the Shapiro-Wilk Test. Yes→ examined outliers and then used paired sample t tests and one sample t tests. No →used nonparametric Wilcoxon Signed-Ranks Test. (2) Outlier: Used standard deviation of 3. Yes → deleted the outliers. No → used all data. Pairwise comparisons: (1) Compared the mean of within-syntactic (Gram) correlations with the mean of across-task correlations. (2) Compared the mean of within-semantic (SCon) correlations with the mean of across-task correlations. (3) Compared the mean of within-syntactic (Gram) correlations with the mean of within-semantic (SCon) correlations. We used Bonferroni correction (p=0.05/3=0.0167) to correct for multiple comparisons. Positivity of the within-task correlations: (1) If the mean of within-syntactic (Gram) correlations was significantly greater than 0. (2) If the mean of within-semantic (SCon) correlations was significantly greater than 0.	The logic for outcome interpretations is the same for all four regions. There are 4 interpretations based on all possible outcomes as described below (also see Figure 2 for illustration). (1) Strong evidence for syntactic specialization: within-syntactic correlations > 0 and within-syntactic correlations > across-task correlations. In addition, within-semantic correlations = 0, and within-semantic correlations = across-task correlations. Alternative patterns that also provided some support for syntactic specialization: within-syntactic correlations > 0 and within-syntactic correlations > across-task correlations plus one of the following semantic results: (a) within-semantic correlations > 0, and within-semantic correlations = across-task correlations. This indicated moderate evidence. (b) within-semantic correlations > 0 and within-semantic correlations > across-task correlations, within-syntactic correlations > within-semantic correlations. This indicated weak evidence. (2) Strong evidence for semantic specialization: within-semantic correlations > 0 and within-semantic correlations > across-task correlations. In addition, within-syntactic correlations = 0, and within-syntactic correlations = across-task correlations. Alternative patterns that also provided some support for semantic specialization: within-semantic correlations > 0 and within-semantic correlations > across-task correlations plus one of the following syntactic results: (a) within-syntactic correlations > 0, and within-syntactic correlations = across-task correlations. This indicated moderate evidence. (b) when within-syntactic correlations > 0 and within-syntactic correlations > across-task correlations, within-semantic correlations > within-syntactic correlations. This indicated weak evidence. (3) Evidence for sensitivity to both syntactic and semantic processing: within-syntactic correlations > 0 and within-syntactic correlations > across task correlations, within-semantic correlations > 0 and within-semantic correlations > across-task correlations, and within-syntactic correlations = within-semantic correlations. (4) No evidence for sensitivity to either semantic or syntactic processing: within-syntactic correlations = across-
2. Is the superior temporal gyrus (STG) an integration region sensitive to both semantic and syntactic processing?	The within-syntactic correlations > 0 and the within-syntactic correlations > across task correlations. The within-semantic correlations > 0 and the within-semantic correlations > across-task correlations. In addition, the within-syntactic correlations = the within-semantic correlations.	We re-calculated the paired sample t tests for both semantic and syntactic effects in STG from the Wang et al. (2020) paper on 5-6-year-old children, respectively. This resulted in t(29)=3.382, p=0.002, effect size=0.628 for Gram vs. Across, and t(29)=5.987, p>0.001, effect size=1.112 for SCon vs. Across. In G*Power, for Gram vs. Across, with Tails=1, Effect size dz=0.628, α err prob=0.02, Power=0.90, the needed sample size is 31. For SCon vs. Across, with Effect size=1.112, α err prob=0.02, Power=0.90, the needed sample size is 12.		
3. Is the opercular part of inferior frontal gyrus (IFG.oper) specialized for syntactic processing?	The within-syntactic correlations > 0 and the within-syntactic correlations > across-task correlations. In addition, the within-semantic	We calculated power based on the only study (Newman et al., 2003) that used the same ROI and similar contrasts. They found a double dissociation in the frontal lobe with an effect in IFG.oper for syntactic violated (noun-verb)		

Question	Hypothesis	Sampling plan (e.g., power analysis)	Analysis Plan	Interpretations given different outcomes
	correlations = 0, and the within-semantic correlations = across-task correlations.	sentences versus fixation baseline [F(1,12)=4.10, p>0.07]. Based on this, we calculated the paired sample t(12)=2.02, and effect size =0.584. In G*Power, with Tails=1, Effect size dz=0.584, α err prob=0.02, Power=0.90, the needed sample size is 35.		task correlations and within-semantic correlations = across-task correlations.
4. Is the triangular part of inferior frontal gyrus (IFG.tri) specialized for semantic processing?	The within-semantic correlations > 0 and the within-semantic correlations > across-task correlations. In addition, the within-syntactic correlations = 0, and the within-syntactic correlations = across-task correlations.	We calculated power based on the only study (Newman et al., 2003) that used the same ROI and similar contrasts. They found a double dissociation in the frontal lobe with an effect in IFG.tri for semantic violated (extraneous verb) sentences versus fixation baseline [F(1,12)=5.63, p>0.04]. Based on this, we calculated the paired sample t(12)= 2.373, and effect size=0.685. In G*Power, with Tails=1, Effect size dz=0.685, α err prob=0.02, Power=0.90, the needed sample size is 26.		

Table 2.

Auditory stimuli conditions and examples

Task	Condition	Response	Brief Explanation	Example
Syntactic Task	Gram	Yes	Grammatical	Every day, they play one game
	FVio	No	Finiteness violation	He dropping one book
	PVio	No	Plurality violation	She is fixing two clock
	PC	Yes	Perceptual control	"Sh – Sh"
Semantic Task	SCon	Yes	Strongly congruent	Last week, she baked two cakes
	WCon	Yes	Weakly congruent	He does not break two glasses
	InCon	No	Incongruent	They are bouncing one paper
	PC	Yes	Perceptual control	"Sh – Sh"

Table 3.

Descriptive statistics for the in-scanner task performance

Task	Condition	Accuracy (%)		Reaction Time (ms)	
		Mean (SD)	Range	Mean (SD)	Range
Syntactic Task	Gram	83.7 (9.8)	60-100	2335 (284)	1819-3055
	FVio	72.4 (19.3)	20-100	3661 (297)	2836-4320
	PVio	89.0 (8.5)	65-100	2281 (212)	1849-2910
	PC	96.3 (5.1)	75-100	1574 (735)	665-4175
Semantic Task	SCon	88.3 (8.5)	70-100	1807 (278)	1276-2480
	WCon	79.0 (13.2)	45-100	1920 (248)	1421-2600
	InCon	90.8 (8.1)	70-100	1915 (260)	1321-2499
	PC	97.8 (3.6)	90-100	1807 (278)	611-3520

Table 4.

The univariate voxel-wise results for each task and task comparisons within the combined language mask.

Type	Brain regions	Brodmann Area	Peak (MNI) Coordinate	Number of voxels	T value
Correct sentences	Gram > PC in the Syntactic Task				
	Left superior + middle temporal gyrus	22/21	−62 −10 4	2525	17.50
	SCon > PC in the Semantic Task				
	Left inferior frontal gyrus -triangular	45	−46 26 −2	56	3.99
	Left superior + middle temporal gyrus	22/21	−62 −12 4	2649	18.15
	Left inferior frontal gyrus -triangular	45	−52 28 4	132	4.36
	Left middle temporal gyrus	21	−54 −62 2	49	4.06
	Syntactic Task (Gram minus PC) > Semantic Task (SCon minus PC)				
	n.s.				
	Semantic Task (SCon minus PC) > Syntactic Task (Gram minus PC)				
	Left middle temporal gyrus	21	−64 −62 12	31	4.26
Incorrect sentences	FVio > PC in the Syntactic Task				
	Left superior + middle temporal gyrus	22/21	−62 −10 4	2833	16.29
	Left inferior frontal gyrus -triangular	45	−46 26 0	298	6.36
	Left inferior frontal gyrus -opercular	44	−50 12 22	28	4.08
	InCon > PC in the Semantic Task				
	Left superior + middle temporal gyrus	22/21	−62 −10 2	1930	18.01
	Syntactic Task (FVio minus PC) > Semantic Task (InCon minus PC)				
	Left middle + superior temporal gyrus	21/22	−52 −36 8	1145	6.09
	Left inferior frontal gyrus -triangular	45	−54 40 −2	447	5.18
	Left inferior frontal gyrus -triangular	45	−36 26 8	33	4.25
	Left inferior frontal gyrus -opercular	44	−44 10 26	71	4.17
	Left inferior frontal gyrus -triangular	45	−58 24 10	69	4.13
	Semantic Task (InCon minus PC) > Syntactic Task (FVio minus PC)				
	n.s.				

Note: Group maps thresholded at voxel-wise $p < 0.001$ uncorrected (T values > 3.2) and cluster-wise $p < 0.05$ corrected within the combined language mask (the left STG, MTG, and the triangular and opercular part of the left IFG). Clusters with size greater than 25 voxels are reported. Gram = Grammatically correct condition, SCon = Strongly congruent condition, FVio = Finiteness violated condition, InCon = Incongruent condition, PC = Perceptual control condition. n.s. = no significant clusters.