scientific reports

OPEN

Distinct functional connectivity patterns during naturalistic learning by adolescent first versus second language speakers

Sibylla Leon Guerrero[1], Laura Mesite[2] & Gigi Luk[3]✉

Spoken lessons (lectures) are commonly used in schools as a medium for conveying educational content. In adolescence, experience-expectant maturation of language and cognitive systems supports learning; however, little is known about whether or how learners' language experiences interact with this integration process during learning. We examined functional connectivity using fMRI in 38 Spanish–English bilingual (L1-Spanish) and English monolingual (L1-English) adolescents during a naturalistic science video lesson in English. Seed analyses including the left inferior frontal gyrus (pars opercularis) and posterior middle temporal gyrus showed that L1-Spanish adolescents, when learning in their second language (L2), displayed widespread bilateral functional connectivity throughout the cortex while L1-English adolescents displayed mostly left-lateralized connectivity with core language regions over the course of the science lesson. Furthermore, we identified functional seed connectivity associated with better learning outcomes for adolescents with diverse language backgrounds. Importantly, functional connectivity patterns in L1-Spanish adolescents while learning in English also correlate with their Spanish cloze reading. Findings suggest that functional networks associated with higher-order language processing and cognitive control are differentially engaged for L1 vs. L2 speakers while learning new information through spoken language.

Keywords Bilingualism, Functional connectivity, Adolescence, Language comprehension, fMRI, Naturalistic tasks

From kindergarten to college, much of formal instruction relies on spoken language to impart novel information, concepts, and skills. Learners' abilities to acquire new knowledge through spoken language are shaped by neural plasticity arising from both developmental and experiential processes[1]. In development, the maturation of language and cognitive systems, particularly in prefrontal brain regions, supports the integration of increasingly complex knowledge representations, while interacting with experience-dependent mechanisms such as bilingualism. However, little is known about how bilingualism, a context- and experience-dependent mechanism, impacts the learning experience of adolescents being taught in a second language (L2). The current study investigated the functional connectivity of higher-order language and cognitive control regions among bilingual adolescents who spoke Spanish as their first, home language and English as the second, instructional language used in school. Bilingual adolescents were compared with functionally monolingual adolescents who spoke English at both home and in school while they watched a science video lesson presented in English.

Globally, it is common for adolescents to receive expository lessons in school settings where instruction is in the non-dominant language. However, little is known about naturalistic learning in an L2 and how bilinguals draw upon their linguistic and cognitive resources when engaged in expository lessons. Because language processing does not occur in isolation in the brain, understanding how brain regions collectively support learning in naturalistic lessons is one way of gaining insight into the affordances and challenges of expository lessons for L2 learners. Yet, given the paucity of prior research using naturalistic stimuli with bilingual populations, and much less with adolescent bilingual participants, little is known about the neural correlates of language processing in understanding naturalistic discourse.

While oral expository lessons are a significant component of K-12 instruction, most previous research has focused on oral narratives[2–4] and reading comprehension[5–7]. To date, no fMRI research has investigated neural

[1]University of California, Irvine, USA. [2]Harvard Graduate School of Education, Cambridge, USA. [3]McGill University, Montreal, Canada. ✉email: gigi.luk@mcgill.ca

responses to complex oral expository lessons in relation to learners' language backgrounds. Thus, the present study evaluates functional regions supporting learning in adolescents for whom the instructional language, namely English, is their second language (L2), in contrast with those for whom the instructional language is their first language (L1).

Current understanding of the neural organization of language processing is grounded in research with monolingual or native speakers exposed to controlled stimuli in their L1[8], displaying a core network for language processing that is primarily left-lateralized in the temporal and posterior frontal cortex[9]. More specific to the comprehension of complex discourse, left-lateralized regions such as the posterior inferior frontal gyrus (BA44, pars opercularis), central inferior frontal gyrus (BA45, pars triangularis), middle temporal gyrus, superior temporal gyrus, angular gyrus, anterior temporal lobe (temporal pole) and insula are important for higher-order meaning composition engaging both semantic and syntactic processing[10].

For bilingual adult speakers, both word-level semantic and higher-order language networks appear to largely overlap across languages in bilinguals[11,12]. The most common differences in word and sentence level processing across bilinguals' L1 and L2 have been found in the magnitude rather than the distribution of activation, with bilinguals displaying differential levels of activation in the same core language regions[13–15]. In addition, bilinguals have been found to display stronger bilateral activation of homologous regions when compared to monolinguals, particularly the inferior frontal and medial temporal lobes[13,14]. Although fewer studies have investigated more complex, connected discourse, similar patterns are also seen with audio narratives in the L1 and L2, with functional activation in anterior temporal regions, weaker, more variable and/or less left-lateralized when listening in the L2[16,17] and with activation patterns moderated by L2 proficiency[18].

Beyond core language regions, the engagement of cognitive control regions can also differ in bilingual, as compared to monolingual, speakers, depending on language proficiency, experimental task, and language context[14,19,20]. At the word and sentence level, the anterior cingulate (associated with context monitoring in language comprehension) and basal ganglia (associated with response selection in language production) are the most common control regions found to display activation differences in bilingual vs. monolingual processing[14,20–22].

While prior research on bilingual discourse comprehension has focused on narrative processing in adults, the current study investigated the processing of expository discourse in adolescent monolingual L1 and bilingual L2 learners under a relatively naturalistic learning scenario similar to what they encounter in school. We focus on adolescents, who potentially experience both developmental and maturational changes in brain structure and function as well as changes stimulated by their bilingual experience after being educated in L2 (English) for a period. In the study, we ask, (1) what are the functional connectivity networks associated with connected discourse processing in L1 English-speaking adolescents and in L1 Spanish-speaking adolescents when learning new information in a naturalistic science lesson in English; and (2) whether and how patterns of neural engagement during learning might be associated with individual differences in L1 and L2 proficiency as measured by standardized language assessments.

To address these questions, we first examined whole-brain functional connectivity during an Earth science video lesson for five a priori regions (L-IFG, L-MTG, L-PCG, L-SMA, and L-INS) that have been demonstrated to play a role in higher-order meaning composition and have been associated with bilingual experience in Activation Likelihood Analysis (ALE) meta-analyses of higher order semantic and syntactic processing[10,23]. Seed-based connectivity analysis provides insight into the ways in which such priori regions-of-interest (ROIs) might be functionally associated with other anatomical brain regions. We thus used seed analysis (AFNI 3dGroupIncorr[24,25]) with ROIs derived from the adult literature due to the limited research on adolescents to examine group commonalities and differences in functional connectivity with 19 Spanish–English and 19 English-only speakers. We further assessed the correlations between individual differences in participants' comprehension of the science lesson video, as well as their L1 and (for Spanish speakers) L2 language proficiency with seed connectivity while controlling for age, nonverbal intelligence, and parental education, factors which have been demonstrated in prior literature to impact functional networks involved in language and higher order processing[26–28]. Extending prior results on L2 language processing in children[29,30] and adults[31,32], we hypothesized that: compared to functionally monolingual native L1 English speakers, bilingual L1 Spanish/L2 English speakers would display wider functional connectivity for regions associated with the left IFG and left MTG seeds with bilateral homologues and with control regions such as the anterior cingulate cortex and basal ganglia when watching an educational video in English and that more extensive functional connectivity would be associated with individual differences in L1 and L2 proficiency.

Results

Behavioral results

Behavioral results are summarized in Table 1. Participants' recall and comprehension of the science video averaged 72% (m = 0.72, sd = 0.18) on comprehension questions about the video content overall, while L1 English speakers scored on average 15 percentage points higher than L1 Spanish speakers (Wilcoxon $p = 0.01$). These video comprehension results were consistent with L1 English speakers overall stronger English language skills as measured in standardized tasks: on average, L1 English speakers displayed higher scores than L1 Spanish speakers on the English vocabulary measure (Wilcoxon $p < 0.001$) and English cloze comprehension task (Wilcoxon $p < 0.001$). Spanish expressive vocabulary (m = 86.7, sd = 13.4) and Spanish cloze comprehension task (m = 88.6, sd = 14.1) scores for L1 Spanish speakers fell on average within one standard deviation of the population mean. In addition, L1 English speakers scored on average two-thirds of a standard deviation higher than L1 Spanish speakers (Wilcoxon $p = 0.002$) on the nonverbal reasoning measure (see Table 1), which was administered in the participant's preferred language, English or Spanish. This group difference in reasoning scores was largely due

	Combined (n = 38) m (sd)	L1 English (n = 19) m (sd)	L1 Spanish (n = 19) m (sd)	Group difference (Wilcoxon *p*)
Video comprehension[a]	0.72 (0.18)	0.80 (0.14)	0.64 (0.19)	0.15 (0.01)
Nonverbal[b] reasoning	111.5 (9.9)	116.5 (9.3)	106.5 (8.0)	9.9 (0.002)
English vocabulary[b]	99.1 (26.6)	117.1 (9.1)	81.1 (26.1)	36.1 (< 0.001)
Spanish vocabulary[b]	–	–	86.7 (13.4)	–
English cloze comprehension[b]	96.9 (22.6)	110.3 (15.0)	83.5 (21.1)	26.8 (< 0.001)
Spanish cloze comprehension[b]	–	–	88.6 (14.1)	–

Table 1. Language and cognitive characteristics. [a]Proportion correct; group difference after rounding. [b]Standard scores with M = 100, SD = 15.

to three particularly high-scoring L1 English speakers as on average, both L1 English and L1 Spanish speakers scored at or above the population mean of 100 in nonverbal reasoning. We retained these high scorers in the sample due to MRI sample size considerations; however, due to this group difference, all subsequent fMRI analyses controlled for nonverbal reasoning scores. Additionally, we conducted a sensitivity analysis for L1 Spanish minus L1 English seed correlations removing these high scorers. While group differences were slightly magnified, as might be expected with unbalanced group sizes, seed correlation results were largely unchanged from the full sample analysis (see Supplementary Table 1).

Video lesson seed correlations

Consistent with processing of multimodal and complex audiovisual stimuli, there was widespread functional connectivity of all seeds with brain regions implicated in language comprehension and audiovisual processing when controlling for age, nonverbal reasoning, and parental education regardless of adolescents' language background (see Supplementary Fig. 1 and Supplementary Table 2). Specifically, extensive clusters were centered in each of the five seed regions, which were identified in prior research with adults as core regions for higher-order language processing (see "Methods", Table 4). There were also significant functional correlations ($p < 0.001$, $\alpha < 0.05$) of seeds with large clusters in bilateral STG, MTG, MFG, SMA, cingulate, and cerebellum, as well as left IFG. See Supplementary Fig. 1 and Supplementary Table 2 for all whole sample and language group seed correlations.

As displayed in Fig. 1 and Table 2, on average, L1 Spanish speakers displayed more extensive seed connectivity than L1 English speakers during the video lesson, when controlling for age, nonverbal IQ and parental education and after cluster correction at $p < 0.001$, $\alpha < 0.05$. Greater recruitment of seed homologues (L1 Spanish – L1 English) was seen with L-MTG (R-MTG, 321vx, MNI 67, – 23, 1) and L-INS (R-INS, 1367vx, MNI 33, 23, 7). This pattern of greater right hemisphere recruitment also extended beyond homologues to language-related (L-MTG seed to R-STG/MTG, 321vx, MNI 67, – 23), visual processing (L-INS to R-cuneus, 560vx, MNI 11, – 81, 46) and cognitive control-related (L-SMA to R-INS, 1367vx, MNI 33,23,7) brain regions. There were no clusters where positive or negative correlated functional connectivity for the L1 English speakers was greater than for the L1 Spanish group after cluster correction at $p < 0.001$, $\alpha < 0.05$.

Video comprehension

For the whole group, seen in Fig. 2a and Table 3a, individual differences in video lesson comprehension were associated with positively correlated activity in language processing regions and negatively correlated activity in higher-order visual processing regions. Specifically, in language processing regions, higher levels of video comprehension were associated with lower functional correlations in the L-IFG seed cluster ($r = -0.28$,

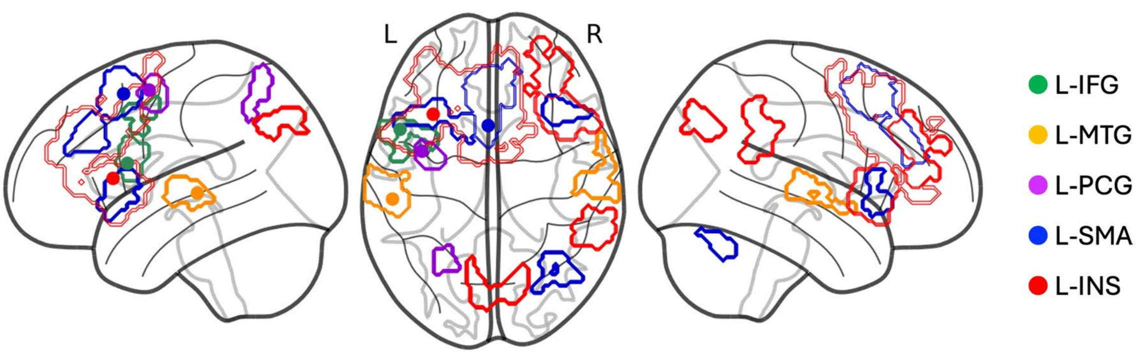

Figure 1. L1 Spanish minus L1 English voxel-wise seed connectivity ($p < 0.001$, $\alpha < 0.05$) for: (**a**) L-IFG, (**b**) L-MTG, (**c**) L-PCG, (**d**) L-SMA and (**e**) L-INS seeds. Cluster outlines display positive cluster correlations; filled clusters indicate negative cluster correlations. Only statistically significant clusters are shown.

Seed (MNI) & correlated cluster peak	Voxels	Sign	X	Y	Z
(1) L-IFG, BA 44 (−50, 12, 16)					
L IFG	496	+	−51	11	15
L PCG	180	+	−49	3	45
(2) L-middle MTG (−54, −26, 0)					
L MTG	485	+	−55	−27	−1
R STG to Heschl's gyrus & MTG	321	+	67	−23	1
(3) L-PCG (−38, 0, 56)					
L PCG	282	+	−39	−1	55
L superior parietal lobule	245	+	−25	−59	57
(4) L-SMA (−2, 14, 54)					
L SMA	1386	+	−3	13	53
L insula to caudate	518	+	−39	19	3
R insula	294	+	33	23	−5
R cerebellum, crus 1	174	+	41	−63	−29
(5) L-INS (−32, 20, 8)					
L insula	7198	+	−33	19	7
R insula	1367	+	33	23	7
R middle frontal gyrus	729	+	31	41	33
R supramarginal gyrus	579	+	55	−39	31
R cuneus	560	+	11	−81	46

Table 2. L1 Spanish minus L1 English: Voxel-wise seed connectivity ($p < 0.001$, $\alpha < 0.05$) for L-IFG, L-MTG, L-PCG, L-SMA and L-INS seeds, with cluster peak coordinates (MNI).

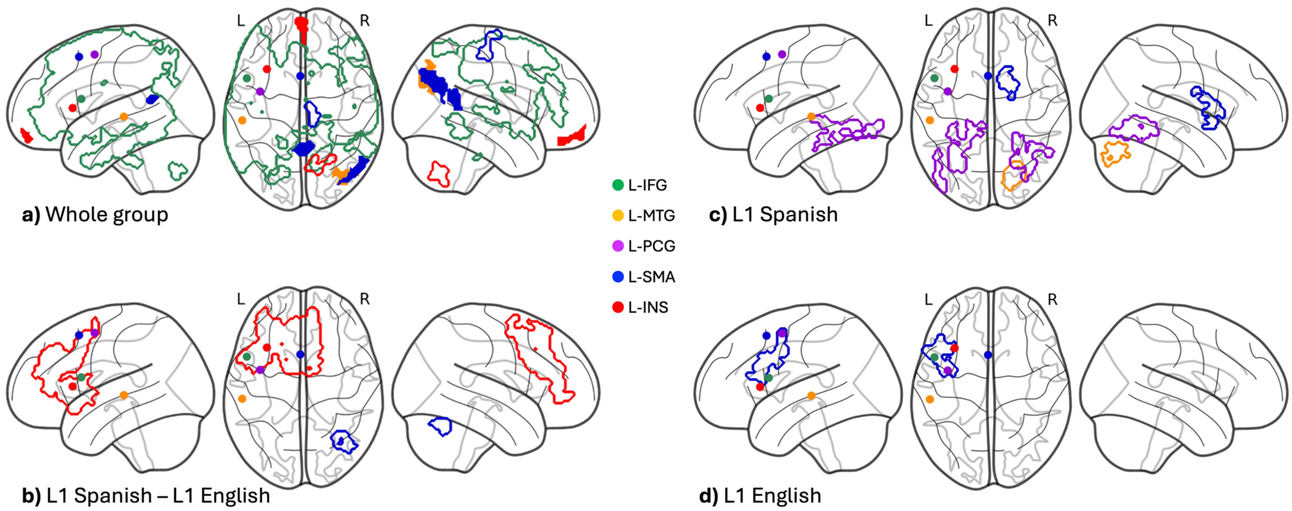

Figure 2. Correlation of video lesson comprehension with voxel-wise seed connectivity clusters ($p < 0.001$, $\alpha < 0.05$) for (**a**) Whole group (n = 38); (**b**) L1 Spanish minus L1 English; (**c**) L1 Spanish only (n = 19); and (**d**) L1 English only (n = 19). Cluster outlines display positive cluster correlations; filled clusters indicate negative cluster correlations. Only statistically significant clusters are shown.

bootstrapped 95% CI [−0.53, −0.03]), as well as lower positively correlated connectivity between L-IFG and R-MTG ($r = -0.38$, bootstrapped 95% CI [−0.65, −0.02]), R-precuneus ($r = -0.26$, bootstrapped 95% CI [−0.51, −0.00]), and L-cerebellum VII ($r = -0.46$, bootstrapped 95% CI [−0.70, −0.15]). In visual processing regions, better video comprehension was associated with higher negatively correlated functional connectivity, such as for the L-SMA seed to bilateral precuneus/PCC (308vx, $r = 0.3$, bootstrapped 95% CI [0.03, 0.63]).

Video comprehension was also associated with group differences (L1 Spanish minus L1 English) in seed connectivity (Fig. 2b and Table 3b). While L1 English speakers on average displayed smaller seed correlations than L1 Spanish speakers of L-SMA to R-cerebellum, this connectivity was associated with better video comprehension in L1 English speakers (178vx, $r = 0.47$, bootstrapped 95% CI [0.14, 0.74]) but not in L1 Spanish speakers. In addition, L1 Spanish speakers on average displayed greater functional connectivity than L1 English speakers in an extensive cluster surrounding the L-INS seed (6927vx), and for L1 Spanish speakers only, higher

Seed (MNI) & correlated cluster peak	Voxels	Sign	X	Y	Z	*r*	95% CI
(a) Whole group							
(1) L-IFG, BA 44 (− 50, 12, 16)							
L IFG	39,210	+	− 51	11	15	− 0.28	[− 0.53, − 0.03]
R MTG	378	+	59	− 53	− 3	− 0.38	[− 0.65, − 0.02]
R precuneus	246	+	11	− 69	61	− 0.26	[− 0.51, − 0.00]
L cerebellum VII	189	+	− 25	− 75	− 55	− 0.46	[− 0.7, − 0.15]
(2) L-MTG (− 54, − 26, 0)							
R mid-occipital/angular gyrus	279	−	41	− 73	33	0.3	[0.06, 0.53]
(4) L-SMA (− 2, 14, 54)							
R angular gyrus	358	−	49	− 75	35	0.29	[0.02, 0.54]
R/L precuneus	308	−	3	− 51	15	0.3	[0.03, 0.63]
R SMA	155	+	7	− 15	73	0.3	[0.08, 0.51]
(5) Left insula (− 32, 20, 8)							
L/R rectal gyrus	189	−	− 1	63	− 19	0.32	[0.01, 0.55]
R cerebellum VIII	188	+	15	− 71	− 55	0.36	[0.09, 0.59]
(b) L1 Spanish minus L1 English							
(4) L-SMA (− 2, 14, 54)							
R cerebellum (L1 Eng. only)	178	+	41	− 63	− 29	0.47	[0.14, 0.74]
(5) Left insula (− 32, 20, 8)							
L insula (L1 Spa. only)	6927	+	− 33	19	7	0.43	[0.04, 0.75]
(c) L1 Spanish only							
(2) L-MTG (− 54, − 26, 0)							
R cerebellum, crus 2	326	+	17	− 83	− 43	− 0.42	[− 0.69, − 0.07]
(3) L-PCG (− 38, 0, 56)							
L fusiform gyrus	673	+	− 35	− 61	− 15	0.55	[0.12, 0.77]
R fusiform gyrus	561	+	29	− 59	− 7	0.53	[0.18, 0.77]
L inf. occipital gyrus	175	+	− 49	− 61	− 15	0.65	[0.36, 0.84]
(4) L-SMA (− 2, 14, 54)							
R caudate	319	+	13	17	5	0.44	[0.12, 0.65]
(d) L1 English only							
(4) L-SMA (− 2, 14, 54)							
L precentral gyrus	785	+	− 43	1	41	− 0.56	[− 0.76, − 0.24]

Table 3. Correlation of video lesson comprehension with voxel-wise seed connectivity clusters ($p < 0.001$, $\alpha < 0.05$) with cluster peak coordinates.

connectivity in this cluster was associated with better video comprehension ($r = 0.43$, bootstrapped 95% CI [0.04, 0.75]). There were no clusters associated with video comprehension scores where positive or negative correlated functional connectivity for the L1 English speakers was greater than for the L1 Spanish group after cluster correction at $p < 0.001$, $\alpha < 0.05$.

The widespread connectivity seen for the LIFG seed in the whole group analysis was not evident when language groups were examined separately. Rather, the most extensive functional connectivity associated with video comprehension in L1 Spanish speakers (Fig. 2c and Table 3c) was between the L-PCG seed and bilateral visual processing regions (L fusiform, 673vx, MNI − 35, − 61, − 15; $r = 0.55$, 95% CI [0.12, 0.77] and R fusiform, 561vx, MNI 29, − 59, − 7; $r = 0.53$, 95% CI [0.18, 0.77]). In L1 English speakers (Fig. 2d and Table 3d), only connectivity between the L-SMA seed and L-PCG was correlated with video comprehension (785vx, MNI − 43, 1, 41; $r = -0.56$, 95% CI [− 0.76, − 0.24]).

English language comprehension

For the whole group (Fig. 3a and Table 4a), English higher order comprehension skills as measured by the English cloze task were positively correlated with seed connectivity between the right cerebellum and the L-SMA seed (2858vx, MNI 9, − 77, − 2; $r = 0.27$, bootstrapped 95% CI [0.01, 0.50]) and from the L-INS seed (294vx, MNI 15, − 71, − 55; $r = 0.40$, bootstrapped 95% CI [0.19, 0.64]). In contrast, functional activity in the R-cuneus (2817vx, MNI 3, − 51, 23) was negatively correlated with activity in the L-IFG seed, with the magnitude of this relationship positively associated with English comprehension ($r = 0.38$, bootstrapped 95% CI [0.01, 0.69]).

Group differences in functional connectivity (Fig. 3b and Table 4b) were negatively associated with English cloze comprehension only in the L-SMA seed cluster for L1 English speakers (765vx, MNI − 3, 13, 53; $r = -0.45$, bootstrapped 95% CI [− 0.73, − 0.02]) and in connectivity of the L-INS seed to the bilateral anterior cingulate for L1 Spanish speakers (942vx, MNI − 5, 25, 33; $r = 0.50$, bootstrapped 95% CI [0.19, 0.77]). There were no clusters

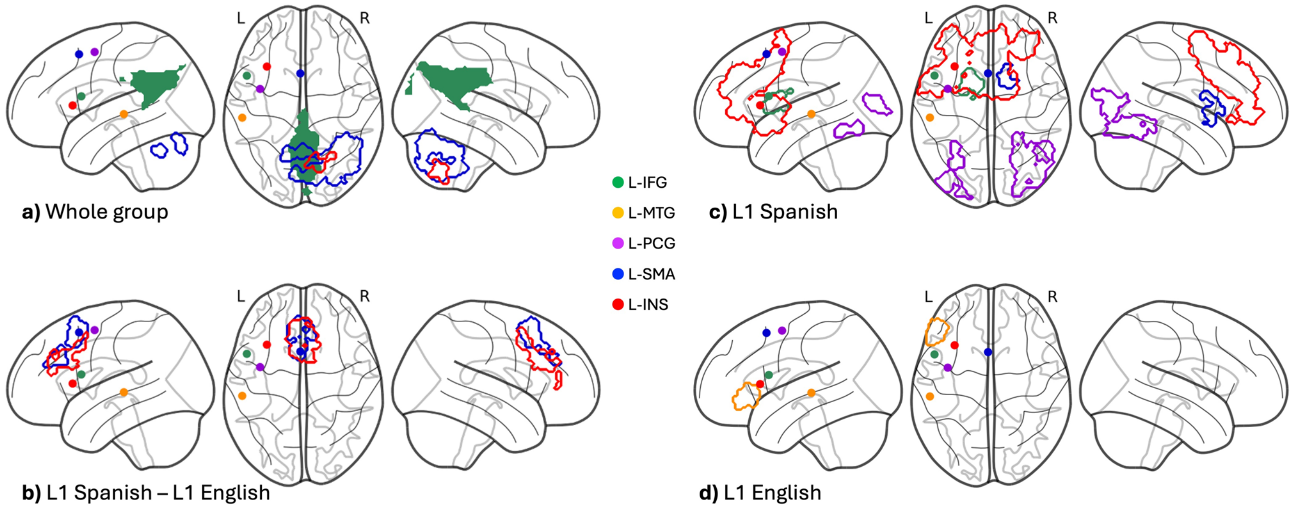

Figure 3. Correlation of English cloze comprehension with voxel-wise seed connectivity clusters ($p < 0.001$, $\alpha < 0.05$) for (**a**) Whole group (n = 38); (**b**) L1 Spanish minus L1 English; (**c**) L1 Spanish only (n = 19); and (**d**) L1 English only (n = 19). Cluster outlines display positive cluster correlations; filled clusters indicate negative cluster correlations. Only statistically significant clusters are shown.

Seed (MNI) & correlated cluster peak	Voxels	Sign	X	Y	Z	*r*	95% CI
(a) Whole group							
(1) L-IFG, BA 44 (− 50, 12, 16)							
R/L-precuneus/PCC	2817	−	3	− 51	23	0.38	[0.01, 0.69]
(4) L-SMA (− 2, 14, 54)							
R-cerebellum, crus 1	2858	+	9	− 77	− 23	0.27	[0.01, 0.5]
(5) L-INS (− 32, 20, 8)							
R-cerebellum, VIII	294	+	15	− 71	− 55	0.4	[0.17, 0.64]
(b) L1 Spanish minus L1 English							
(4) L-SMA (− 2, 14, 54)							
L-SMA (L1 Eng. only)	765	+	− 3	13	53	− 0.45	[− 0.73, − 0.02]
(5) Left insula (− 32, 20, 8)							
L/R mid- to anterior cingulate (L1 Spa. Only)	942	+	− 5	25	33	0.5	[0.19, 0.77]
(c) L1 Spanish only							
(1) L-IFG, BA 44 (− 50, 12, 16)							
L putamen	389	+	− 21	9	5	0.35	[0.02, 0.68]
(3) L-PCG (− 38, 0, 56)							
L precentral gyrus	1304	+	− 39	− 1	55	0.42	[0.07, 0.68]
L inf. parietal lobule	302	+	− 41	− 85	15	0.45	[0.08, 0.73]
L fusiform gyrus	235	+	− 33	− 59	− 13	0.28	[0.02, 0.54]
(4) L-SMA (− 2, 14, 54)							
R caudate	203	+	13	17	5	0.46	[0.09, 0.69]
(5) L-INS (− 32, 20, 8)							
L insula	10,517	+	− 33	19	7	0.37	[0.04, 0.68]
(d) L1 English only							
(2) L-MTG (− 54, − 26, 0)							
L IFG, BA45, pars tri	508	+	− 55	27	− 1	0.51	[0.15, 0.69]

Table 4. Correlation of English cloze comprehension with voxel-wise seed connectivity clusters ($p < .001$, $\alpha < .05$) with cluster peak coordinates.

associated with English cloze comprehension where positively or negatively correlated functional connectivity for the L1 English was greater than for the L1 Spanish group after cluster correction at $p < 0.001$, $\alpha < 0.05$.

In L1 Spanish speakers, qualitative patterns of extensive bilateral functional connectivity are seen in Fig. 3c and Table 4c for the L-PCG seed to occipital visual processing regions and for the L-INS cluster (10517vx, MNI − 33, 19, 7; $r = 0.37$, bootstrapped 95% CI [0.04, 0.68]). In L1 English speakers (Fig. 3d and Table 4d), English cloze comprehension skills were positively associated with connectivity between L-MTG and L-IFG (508vx, MNI − 55, 27, − 1; r = 0.51, 95% CI [0.15, 0.69]).

Spanish language comprehension

For L1 Spanish speakers, we also examined potential associations of Spanish higher-order comprehension skills with functional connectivity while watching the video in L2 English. A separate seed-based analysis including Spanish cloze comprehension as a covariate was conducted only for L1 Spanish speakers. When controlling for age, English vocabulary and English passage comprehension skills, L1 Spanish speakers' Spanish cloze skills were significantly and negatively correlated with functional connectivity in a wide range of clusters associated with the L-IFG and L-SMA seeds (see Fig. 4 and Table 5). For the L-IFG seed, clusters were centered in L-IFG (7489vx, r = − 0.51, bootstrapped 95%, CI [− 0.79, − 0.06]), R-SMG (821vx, r = − 0.46, bootstrapped 95%, CI [− 0.68, − 0.14]), R-SFG (r = − 0.54, bootstrapped 95%, CI [− 0.77, − 0.23]), and L-ACC (180vx, r = − 0.59, bootstrapped 95%, CI [− 0.82, − 0.19]). For the left SMA seed, clusters were centered in bilateral SMA (12,266vx, r = − 0.64, bootstrapped 95%, CI [− 0.81, − 0.23]) but also included right hemisphere regions such as cerebellum, caudate, and angular gyrus (see Table 5).

Discussion

We examined L1 English and L1 Spanish speakers' neural responses to expository oral discourse using a naturalistic and ecologically relevant paradigm in fMRI. Results demonstrate differential patterns of neural recruitment across the two groups. These patterns were reflective of language knowledge in English, for all participants, as well as language knowledge in Spanish, for L1 Spanish, L2 English bilingual participants.

The first research question asked how brain regions associated with connected discourse processing are engaged when adolescents watch and listen to a naturalistic science lesson in either their L1 or L2. L1 Spanish

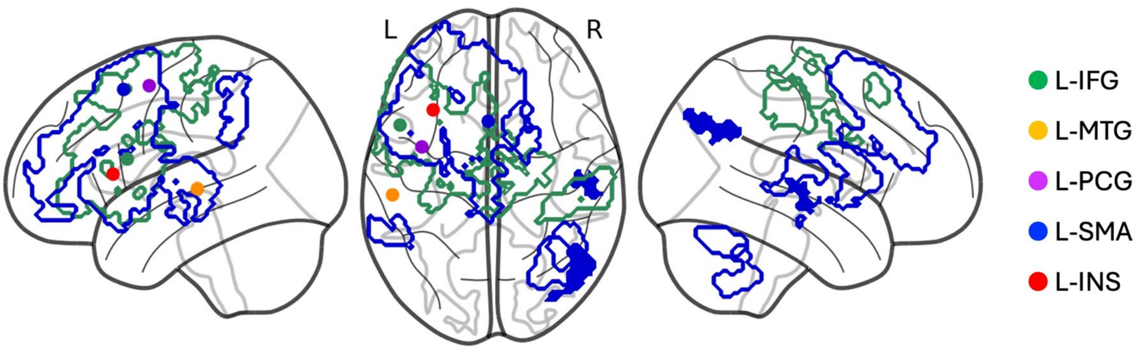

Figure 4. L1 Spanish speakers' correlation of Spanish cloze comprehension with voxel-wise seed connectivity clusters (p < 0.001, α < 0.05). Cluster outlines display positive cluster correlations; filled clusters indicate negative cluster correlations. Only statistically significant clusters are shown.

Seed (MNI) & correlated cluster peak	Voxels	Sign	X	Y	Z	*r*	95% CI
(1) L-IFG, BA 44 (− 50,12,16)							
L IFG	7489	+	− 51	11	15	− 0.51	[− 0.79, − 0.06]
R supramarginal gyrus	821	+	55	− 27	45	− 0.46	[− 0.68, − 0.14]
R sup. frontal gyrus	498	+	27	− 5	61	− 0.54	[− 0.77, − 0.23]
L anterior cingulate	180	+	− 3	1	31	− 0.59	[− 0.82, − 0.19]
(2) L-MTG (− 54, − 26, 0)							
(3) L-PCG (− 38, 0, 56)							
(4) L-SMA (− 2, 14, 54)							
L-SMA to R-SMA	12,266	+	− 3	13	53	− 0.64	[− 0.87, − 0.28]
R cerebellum, crus 1	616	+	41	− 61	− 29	− 0.64	[− 0.88, − 0.27]
R caudate	555	+	15	9	13	− 0.66	[− 0.84, − 0.25]
R angular gyrus	367	–	47	− 75	35	− 0.46	[− 0.68, − 0.11]
L inf. parietal lobule	344	+	− 49	− 47	43	− 0.53	[− 0.77, − 0.18]
R cerebellum VIII	309	+	39	− 57	− 53	− 0.49	[− 0.73, − 0.15]
R sup. temporal gyrus	180	–	53	− 21	− 1	− 0.69	[− 0.86, − 0.41]
(5) Left insula (− 32, 20, 8)							

Table 5. Correlation of Spanish cloze comprehension with voxel-wise seed connectivity clusters (p < 0.001, α < 0.05) (with cluster peak coordinates) among L1 Spanish adolescents.

speakers were hypothesized to display more diffused and heightened functional connectivity than L1 English speakers with bilateral homologues and with control regions such as the anterior cingulate and basal ganglia. Results of the seed-based analysis indicated that in these Spanish–English bilinguals, there was indeed an overall pattern of more extensive connectivity among clusters centered on the functional seeds. L1 Spanish speakers also displayed greater bilateral functional connectivity between a priori seeds to whole brain clusters in homologues and control regions. However, this pattern varied by seed region. For example, compared to L1 English speakers, L1 Spanish speakers displayed more extensive correlated activity between the L-IFG seed with L-PCG, but not with homologous syntax- or language-related regions in the right IFG. This result both contrasts with and complements findings from word- and sentence-level studies of bilingual language processing, which have more broadly found homologous activation for the L-IFG in L2 processing[30,33,34]. The L-MTG seed displayed greater connectivity with right hemisphere homologues of auditory processing regions in the STG and Heschl's gyrus. However, the most extensive differences between L1 Spanish and L1 English speakers were found in the functional connectivity of L-SMA and L-INS seeds, with left and right hemisphere clusters, including homologues, related to both language and visual processing, as well as cognitive control.

Overall, comprehension of the video lesson was supported by connectivity among core language processing regions, such as the L-IFG, L-MTG, and L-SMA and by connectivity of seeds with brain regions involved in more automatized, rule-based language processing, such as the right cerebellum. Further, while we found extensive connectivity between seeds and both language processing as well as visual processing regions, this connectivity appeared to be a hallmark of effortful processing and was associated with lower comprehension of the science lesson.

The second research question asked whether patterns of neural processing were associated with individual differences in L1 and L2 higher-order language skills, as measured by cloze comprehension tasks in English and Spanish. For all adolescents, connectivity of the L-SMA and L-INS seeds with R-cerebellum during the video lesson was positively correlated with English cloze skills and indeed, better video comprehension. This relationship may reflect rule-based lexical and syntactic knowledge engaged while learning from the video lesson. English language skills also appear to ease cognitive control demands that are signaled by the functional connectivity of L-IFG with R-precuneus/PCC during the video lesson. However, drawing on these higher-order language skills may require the engagement of language control during learning for L2 learners, as suggested by the positive relationship between English cloze skills and L1 Spanish speakers' engagement of bilateral mid-cingulate during the video lesson.

Further, our findings suggest that these bilingual adolescents draw upon their L1 (Spanish) language skills during the video lesson, even when the lesson is presented in their L2 (English). For L1 Spanish speakers, connectivity of the core compositional regions of the L-IFG and L-SMA with bilateral language and visual processing, and cognitive control regions was associated with Spanish comprehension skills, such that better Spanish cloze comprehension was associated with lower connectivity in these regions. This inverse association of skilled language with functional connectivity was also evident for video lesson and English cloze comprehension. Recruitment of cognitive control regions is a typical indicator of effortful processing[35]. Bilingual adolescents with stronger Spanish cloze comprehension displayed lower connectivity with control regions, thus suggesting that higher-order comprehension skills make L2 processing easier, regardless of whether those skills were acquired in L2 or L1, i.e., English or Spanish. Indeed, the facilitating effect of L1 higher-order discourse skills was also reflected in an eye-tracking study of a larger sample of adolescents of which the current sample is a subset[36]. In the larger study examining L1 Spanish adolescents' reading outcomes and eye movement behaviors, Spanish cloze skills as indicators of syntactic processing were associated with more efficient online eye movement during reading and better passage comprehension while reading English texts[36].

The current study included 38 participants across two language groups. Consequently, although the results revealed significant outcomes for some brain regions and behavioral correlates, the statistical power to detect group differences in functional connectivity and individual differences in significant behavioral correlates is limited. With functional connectivity analysis and naturalistic task designs, caution should also be used in interpreting the magnitude and sign (e.g. positive vs. negative) of functional correlations as these do not necessarily translate to facilitation or inhibition of cognitive processing. An additional limitation of the current study is the lack of equivalency across language groups on potential confounders such as age, socioeconomic status, and nonverbal IQ. Although the sample groups were representative of student populations found in the U.S., we acknowledge that statistically controlling for these variables in all our analyses does not equate these variables between groups. Finally, among the L1 Spanish speakers, onset age of English acquisition, length of English immersion, and English proficiency were highly positively correlated while these variables were highly negatively associated with Spanish proficiency. Although seed-based models also controlled for English proficiency, we were thus unable to disaggregate within-group heterogeneity in L1 Spanish speakers with respect to these different characteristics of diverse language experience.

The present study demonstrates that even when accounting for L2 vocabulary and L2 comprehension, there are both similarities and differences in how L1 and L2 speakers engage the brain's networks for language and language control. L2 speakers overall display a pattern of engagement of stronger, more widespread, and more bilateral connectivity during learning from lessons in L2, with recruitment of regions associated with procedural skill and cognitive control. Differences emerge particularly in regions associated with higher-order language processing and cognitive control. This finding is especially relevant to adolescents learning academic lessons in their L2 as such lessons often focus on L2 vocabulary development[37,38]. The current findings provide support for the idea that, beyond L2 vocabulary, students draw upon higher-order comprehension skills developed in their L1 to integrate L2 word meanings in understanding L2 discourse. This finding highlights the potential benefits of enhancing comprehension using syntactic and integrative skills in the L1 as a linguistic resource. Further, findings provide evidence of how spoken science lessons in a developing L2 are cognitively demanding, drawing on multiple control regions and processes and highlight the importance of reducing cognitive load when instruction occurs with oral lessons in the L2. This study complements prior adult research on bilingual processing of naturalistic narrative stimuli and calls for further research to better understand with greater precision both the conceptual and linguistic features of the lesson and the individual differences in first and second language proficiency which contribute to L2 comprehension during school learning.

Methods

Sample

Forty-three middle-school children (grade 6, n = 14; grade 7; n = 10, grade 8, n = 13) residing in a metropolitan area in the northeast U.S. were recruited as part of a larger study through community announcements, social media, and outreach to schools. All study procedures including informed consent and participant compensation were approved by the Committee on the Use of Human Subjects at Harvard University in accordance to the ethical principles set forth in the Belmont Report. Parental consent and youth assent were obtained prior to all study procedures. The 43 children had no diagnosis of language, reading, hearing or uncorrected vision difficulty. Of these, 22 spoke English as a first language (L1) and had no extensive or persistent second language immersion experience. These L1 English speakers had minimal proficiency in a non-English language as reported by their parents, likely due to world language instruction at school. The remaining 21 children were L1 Spanish/L2 English bilingual speakers. Of these participants from the two language groups, three L1 English and two L1 Spanish speakers were excluded from the analysis due to excessive head motion (average relative motion > 0.2 mm). After the exclusion of these participants, average head motion did not differ between the two language groups (Kruskal–Wallis $\chi^2 = 0.06$, $p = 0.80$). The final analytic sample included 19 L1 English and 19 L1 Spanish speakers.

The sample's demographic characteristics are presented in Table 6. Participants were 22 female and 16 male children ranging between 11 and 15.2 years of age (m = 12.9, sd = 1.1) with a median grade level of 7th grade. Females and males were evenly distributed in the two language groups. L1 Spanish speakers were on average 1.1 years older (Kruskal–Wallis χ^2 $p = 0.004$) than L1 English speakers but the two groups did not differ significantly in grade level (Fisher's Exact, $p = 0.08$). All mothers of L1 English speakers (n = 19) had completed university at the bachelor's level or higher. Eleven mothers of L1 Spanish speakers had completed university degrees while six had post-secondary education but had not completed a tertiary degree, and two had completed secondary education but had no post-secondary experience. Language groups thus differed in levels of maternal and highest parental education (Fisher's Exact, $p = 0.003$). Because socioeconomic status, particularly parental education, has been shown to impact language development[27], all analyses in this study controlled for parental education as well as participants' age.

L1 Spanish speakers had extensive and ongoing bilingual and biliterate experience as indicated by parent reports on a questionnaire adapted from a previous study[39]. Fourteen of the 19 children had moved to the U.S. from Spanish-speaking countries at varying ages, including Mexico (n = 5), Honduras (n = 2), Colombia (n = 2), El Salvador (n = 1), Nicaragua (n = 1), Peru (n = 1), Venezuela (n = 1) and Spain (n = 1) where Spanish was the dominant language in their home, community, and school. The remaining five participants were born into Spanish-speaking homes in the U.S., spoke Spanish as L1 and acquired English upon entering kindergarten where English was the primary language of instruction for three of these children. The other two of these five children attended a bilingual Spanish–English school in addition to speaking Spanish daily at home. Across the groups, the average number of years of English immersion was 5.0 years (sd = 3.6 years) but this experience also ranged widely from 6 months to 10 years, as illustrated in Table 3. All L1 Spanish children spoke both Spanish and English at home (mean daily Spanish use = 48.5%, sd = 15.9%) and most engaged with Spanish-language printed materials (m = 36.3%, sd = 19.4%), music (m = 23.6%, sd = 19.8%) and multimedia (m = 27.2%, sd = 19.8%) at home on a daily basis. A few participants in this group had also received classroom instruction in French (n = 3), Portuguese (n = 2), Chinese (n = 1), Japanese (n = 1) and/or Latin (n = 1) and self-reported minimal speaking ability in these languages.

L1 English speakers had minimal exposure and usage to a L2, almost entirely in the form of classroom instruction in a restricted range of languages that included Spanish (n = 6), Chinese (n = 5), French (n = 3), and Latin (n = 3) and self-reported minimal speaking and comprehension abilities. Participants in this group had no L2 immersion experience and only limited L2 instruction (m = 1.7 years, sd = 1.2 years) that started on average in the middle school years (m = 10.7 years, sd = 1.8) and ranged between 40 min to 5 h per week. Outside of the foreign language classroom, all school instruction for these children was conducted in English, and languages other than English were minimally present in the home with a non-English language heard in the home on average only 1.3% of the day (sd = 2.7%).

	Combined (n = 38) m (sd)	**L1 English (n = 19) m (sd)**	**L1 Spanish (n = 19) m (sd)**	**Group difference (*p* of Kruskal–Wallis χ^2)**
Sex (female/male)	22/16	11/8	11/8	Fisher's exact ($p = 1.00$)
Grade in school (median)	7th	7th	8th	Fisher's exact ($p = .08$)
Age (years)	12.9 (1.1)	12.4 (1.0)	13.5 (1.0)	1.1 (0.004)
Parent education SES index	3.7 (0.6)	4.0 (0.0)	3.5 (0.7)	0.5 (0.03)

Table 6. Participant demographic characteristics.

Stimulus

The functional stimulus was an approximately 6-min-long naturalistic video presentation of a lesson on earth science and earthquakes. The video topic, content, and lesson text were constructed by examining middle school state standards and benchmarks for the earth sciences and developing content to match the 7th grade, or median grade level for the sample. In the lesson video, called "Earthquakes," a narrator who was a former middle school science teacher orally relayed the lesson with graphic illustrations and photographs related to lesson concepts used as visual backdrops (see study materials available on OSF, https://osf.io/2yuex/?view_only=611d3d0ade6f4536be66a100a56043c3, for a sample scene and full audio text). The video included 19 distinct visual backdrops illustrating audio content with no animation in the graphic illustrations and photographs. The video stimulus was presented using PsychoPy[40] and projected with an LCD projector. Participants viewed the video with a headcoil-mounted mirror, which captured a rear-projection screen fixed at the end of the scanner bore. Participants listened to the audio using Sensimetrics S14 MRI-compatible insert earphones (Sensimetrics, Ltd.).

Experimental design and procedure

Adolescents completed two testing sessions: the first session included informed consent and behavioral assessment as well as mock scanning to familiarize participants with MRI scanning procedures and lastly, the language and social background questionnaire completed by parents[39]. During the second testing session and before entering the scanner, adolescents were asked to share their knowledge of the video topic with their responses audio recorded. Participants were confirmed to have minimal knowledge about the topic and concepts presented in the video. They next completed an fMRI session that included T1-weighted structural acquisition and a single BOLD run acquired while watching the video lesson about earth science and earthquakes. Before the start of the video run, participants were reminded to pay close attention to the video as they would be asked questions about the content. Immediately after the video, participants were removed from the scanner and asked to freely recall what they remembered from the video as well as respond orally to specific questions about the video content. The scan session included additional scans (resting state, task and diffusion imaging) before the final video scan as part of the larger study. Results from the additional scans are not reported here.

Behavioral and language assessment

Video comprehension

Recall and understanding of the science video was assessed using an experimental measure consisting of one response prompt ("Please tell me everything you remember about the video.") and ten specific questions about video content that included eight questions requiring factual recall (e.g., "What is the inner core made out of?"); and two inferential questions that required integration from the video content and background knowledge (e.g., "Why might it be dangerous to live on or near a fault line?"). The full list of questions can be found on OSF at [https://osf.io/2yuex/?view_only=611d3d0ade6f4536be66a100a56043c3]. All questions were initially presented to all participants in English. For L1 Spanish speakers who had difficulty comprehending or responding to English questions, these questions were repeated in Spanish. Both English and Spanish answers were accepted and included in the scoring. Two trained raters independently scored all responses as incorrect (0 pts.), partially correct (1 pt.) or fully correct (2 pts.) based on a rubric that was co-constructed by a team of researchers with teaching experience. Raters agreed on 90% of item scores, calculated by (total items scored – discrepant scores)/(total items scored) × 100%. For discrepant scores, raters then discussed discrepant responses and came to an agreement on a final score.

Demographic and language experience

Parents of participants completed an adapted Language, Social and Background Questionnaire (LSBQ)[39] about their child's language history and use, as well as the language use of family members interacting with the child. Parents completed the LSBQ in their preferred language (English or Spanish) with the assistance of a research assistant who spoke Spanish for L1 Spanish families. The full LSBQ can be found on OSF at [https://osf.io/2yuex/?view_only=611d3d0ade6f4536be66a100a56043c3].

Nonverbal reasoning

To assess nonverbal reasoning skills, participants completed the Kaufman Brief Intelligence Test, 2nd Edition (KBIT-2)[41], Matrices subtest, a standardized assessment of reasoning ability. Participants completed the nonverbal reasoning portion, Matrices, which requires completing a visual analogy by pointing to the correct image. Participants received instructions and prompts in either Spanish or English according to their preference. The KBIT-2 Matrices subtest yields a raw score of items correct and an age-corrected standard score ($M = 100$, $SD = 15$) and demonstrates adequate internal consistency (Cronbach's alpha = 0.88) and concurrent validity[41].

English/Spanish vocabulary

English vocabulary was assessed with the Woodcock-Muñoz Language Survey, Revised, English (WMLS-R English)[42] with Normative Update[43], Picture Vocabulary subtest. Spanish vocabulary was assessed with a parallel and equated Spanish form of the WMLS-R, English, the Woodcock-Muñoz Language Survey, Revised, Spanish (WMLS-R Spanish)[44] with Normative Update[43], Vocabulario sobre dibujos subtest. The WMLS-R, English and Spanish versions are standardized assessments of English language proficiency with norms for a nationally representative sample of U.S. children updated in 2005. The Spanish form was calibrated with native Spanish speakers from seven countries spanning North, Central and South America, with U.S. subjects representing 7.3% of the calibration sample study[45]. The Picture Vocabulary/Vocabulario sobre dibujos subtest asks individuals to orally identify pictured objects of decreasing word frequency. Each subtest yields a raw score of the number of items correctly answered and an age-scaled standard score (M = 100, SD = 15). Standard scores were used in subsequent analysis. For the ages represented in the current study, internal consistency reliability coefficients for Picture Vocabulary range from 0.91 to 0.94. Validity measures were not published for the age groups and subtests represented in the current study[45].

Connected discourse comprehension skills were measured in English with the Woodcock-Muñoz Language Survey, Revised, English (WMLS-R English)[42] with Normative Update[43], Passage Comprehension subtest and in Spanish with a parallel and equated Spanish form of the WMLS-R, English, the Woodcock-Muñoz Language Survey, Revised, Spanish (WMLS-R Spanish)[44] with Normative Update[43], Comprensión de textos subtest. The Passage Comprehension/Comprensión de textos subtests are cloze tasks asking the participant to read one or more sentences and orally supply a missing word. The task starts with stimuli consisting of a single, very brief sentence, expanding to two sentences and/or a longer sentence stimulus as the difficulty of the task increases. Each subtest yields a raw score, which is the number of items correctly responded and an age-scaled standard score (M = 100, SD = 15); the latter was used in subsequent analysis. For the ages represented in the current study, internal consistency reliability coefficients for WMLS-R Passage Comprehension range from 0.80 to 0.83.

Descriptive statistics

Inspection of the raw data revealed that in most cases, the distribution of behavioral data violated assumptions of normality and variance homogeneity, and in regression models, of sphericity and homoscedasticity. This analysis thus utilized non-parametric tests implemented in R (R Core Team, 2018) to compute basic descriptive correlations and first-level group comparisons. Specifically, the analysis employed Wilcoxon (Mann Whitney) signed rank tests (R package *coin*)[46] and BCa (bias corrected and adjusted) bootstrapped 95% confidence intervals to test differences in sample means (R package *boot*)[47].

MRI acquisition

Participants were scanned in a 3 T Magnetom scanner (Prisma, Siemens) at the Center for Brain Science at Harvard University, using a 32-channel head-neck coil (Siemens). A high-resolution anatomical image was collected using a T1-weighted MPRAGE single-band, multi-echo pulse sequence (176 slices; voxel size = 1 × 1 × 1 mm; FoV = 256 mm; TR = 2260 ms; TE = 1.69 ms; 3.55 ms, 5.41 ms, 7.27 ms; 7° flip angle; GRAPPA with acceleration factor of 4). Functional images were collected using a multi-band[48–50] single-echo echo planar T2* BOLD sequence (64 interleaved slices; voxel size = 2.3 × 2.3 × 2.3 mm; FoV = 207 mm; TR = 650 ms; TE = 34.8 ms; 52° flip angle; multi-band acceleration factor = 8). Pre-scan training, cushioning, and sensory feedback (a tape across the forehead) were used to minimize head movement. In addition, in-ear headphones were used to convey the video's audio track and were covered by over-ear headphones to minimize scanner noise.

fMRI data preprocessing

MRI data were preprocessed using a standard pipeline from FMRIPREP version 20.1.1 (RRID:SCR_016216)[51], a Nipype based tool[52]. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0[53] and skull-stripped using antsBrainExtraction.sh v2.1.0. Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.1[54], and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle[30]. Spatial normalization to the ICBM MNI152 Nonlinear Asymmetrical template version 2006 (MNI152NLin2006)[55] was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0[32], using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white matter (WM) and gray matter (GM) was performed on the brain-extracted T1w using FSL's FAST[56]. Functional data was slice time corrected using 3dTshift from AFNI v16.2.07[24] and motion corrected using mcflirt (FSL v5.0.9). "Fieldmap-less" distortion correction was performed by co-registering the functional image to the same-subject T1w image with intensity inverted[57,58] constrained with an average fieldmap template[59], implemented with antsRegistration (ANTs). This was followed by co-registration to the corresponding T1w using boundary-based registration[39] with nine degrees of freedom, using bbregister (FreeSurfer v6.0.1). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation. ICA-based Automatic Removal Of Motion Artifacts (AROMA) was then used to generate non-aggressively denoised functional data[60] including demeaning, smoothing to 6 mm full-width half-maximum using FSL's SUSAN kernel[61] and normalization to the MNI152NLin2006 template. For more details of the pipeline see https://fmriprep.readthedocs.io/en/sTable/workflows.html.

Denoising for scanner and physiological noise was implemented in xcpEngine[62] and included linear detrending, temporal filtering (forward-and-reverse Butterworth filter, 0.01–0.08 Hz) and subsequent confound regression for white matter and CSF regressors and their squared derivatives. Five participants, three L1 English and two L1 Spanish speakers, were excluded from the analysis due to excessive head motion (average relative motion > 0.2 mm). After the exclusion of these participants, average head motion did not differ between the two language groups (Kruskal–Wallis $\chi^2 = 0.06$, $p = 0.80$).

Seed-based correlation analysis

Five cortical regions of interest were selected as seeds based on two Activation Likelihood Analysis (ALE) meta-analyses of higher order semantic and syntactic processing: Rodd and colleagues[10] included 54 studies of receptive language while Walenski and colleagues[23] analyzed 45 studies, with 37 of them involving complex syntax. Regions which overlapped across the two meta-analyses when examining auditory comprehension and syntax were (1) left IFG (BA44, pars opercularis, MNI −50, 12, 16); (2) left MTG (MNI −54, −26, 0); (3) left precentral gyrus (MNI −38, 0, 56); (4) left SMA (−2, 14, 54); and (5) left insula (−32, 20, 8).

Using these five regions common to the two meta-analyses, we extracted mean BOLD time series within 2 mm spherical seeds from preprocessed and denoised data and computed whole-brain correlations using AFNI's 3dGroupIncorr[24,25]. The resulting voxel-wise correlation maps were Fisher-transformed to z-scores and thresholded at an uncorrected $p = 0.001$ and cluster-corrected to $p \leq 0.05$. Thresholding was calculated using AFNI's 3dClustSim (Monte Carlo simulation) using bi-sided nearest neighbor clustering to define contiguous voxels, with family-wise error rate (FWER) correction for multiple comparisons[63,64]. Group comparisons in 3dGroupIncorr employed two-sample, unpaired t-tests with pooled variance estimators and controlled for participant age, nonverbal reasoning, and maternal education to account for differences in these variables across language groups. English expressive vocabulary was additionally included in seed analysis as a covariate in the group comparison of English cloze comprehension to control for the role of word-level language proficiency. A supplemental within-group seed analysis of L1 Spanish speakers included Spanish passage comprehension as an individual-level covariate. To examine the potential role of L1 sentence integration skills and control for L2 lexical and comprehension skills in neural processing, this supplemental analysis also included English vocabulary and English passage comprehension as control variables along with participant age, nonverbal IQ and parental education. Masks created from resulting t-statistics were plotted with nilearn (RRID:SCR_001362) and scikit-learn[65].

Data availability

The protocol, code, and data that support the findings of this study are available at Open Science Framework at [https://osf.io/2yuex/?view_only=bf89a3c48b7f4afab69699dbc0306168].

Received: 30 April 2024; Accepted: 6 August 2024
Published online: 16 August 2024

References

1. Greenough, W. T., Black, J. E. & Wallace, C. S. Experience and brain development. *Child Dev.* **58**, 539–559 (1987).
2. Bhattasali, S. *et al.* Localising memory retrieval and syntactic composition: An fMRI study of naturalistic language comprehension. *Lang. Cogn. Neurosci.* **34**, 491–510 (2019).
3. Rocca, R. *et al.* Language beyond the language system: Dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI. *NeuroImage* **216**, 116128 (2020).
4. Tang, X. *et al.* Distinct brain state dynamics of native and second language processing during narrative listening in late bilinguals. *NeuroImage* **280**, 120359 (2023).
5. Birba, A. *et al.* Motor-system dynamics during naturalistic reading of action narratives in first and second language. *NeuroImage* **216**, 116820 (2020).
6. Hsu, C.-T., Clariana, R., Schloss, B. & Li, P. Neurocognitive signatures of naturalistic reading of scientific texts: A fixation-related fMRI study. *Sci. Rep.* **9**, 10678 (2019).
7. Wang, K. *et al.* Characterizing and decomposing the neural correlates of individual differences in reading ability among adolescents with task-based fMRI. *Dev. Cogn. Neurosci.* **37**, 100647 (2019).
8. Meyer, L. & Friederici, A. D. Chapter 48—Neural systems underlying the processing of complex sentences. in *Neurobiology of Language* (eds. Hickok, G. & Small, S. L.) 597–606 (Academic Press, 2016). https://doi.org/10.1016/B978-0-12-407794-2.00048-1.
9. Hagoort, P. The core and beyond in the language-ready brain. *Neurosci. Biobehav. Rev.* **81**, 194–204 (2017).
10. Rodd, J. M., Vitello, S., Woollams, A. M. & Adank, P. Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. *Brain Lang.* **141**, 89–102 (2015).
11. Perani, D. & Abutalebi, J. The neural basis of first and second language processing. *Curr. Opin. Neurobiol.* **15**, 202–206 (2005).
12. Van de Putte, E., De Baene, W., Brass, M. & Duyck, W. Neural overlap of L1 and L2 semantic representations in speech: A decoding approach. *NeuroImage* **162**, 106–116 (2017).
13. Del Maschio, N. & Abutalebi, J. Chapter 15. Neurobiology of bilingualism. in *Studies in Bilingualism* (eds. Miller, D., Bayram, F., Rothman, J. & Serratrice, L.) vol. 54 325–346 (John Benjamins Publishing Company, 2018).
14. Sulpizio, S., Del Maschio, N., Fedeli, D. & Abutalebi, J. Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* **108**, 834–853 (2020).
15. Brignoni-Pérez, E., Jamal, N. I. & Eden, G. F. Functional neuroanatomy of English word reading in early bilingual and monolingual adults. *Hum. Brain Mapp.* **43**, 4310–4325 (2022).
16. Dehaene, S. *et al.* Anatomical variability in the cortical representation of first and second language. *NeuroReport* **8**, 3809–3815 (1997).
17. Perani, D. *et al.* Brain processing of native and foreign languages. *NeuroReport-Int. J. Rapid Commun. Res. Neurosci.* **7**, 2439–2444 (1996).
18. Hesling, I., Dilharreguy, B., Bordessoules, M. & Allard, M. The neural processing of second language comprehension modulated by the degree of proficiency: A listening connected speech fMRI study. *Open Neuroimaging J.* **6**, 44–54 (2012).
19. Abutalebi, J. & Green, D. Bilingual language production: The neurocognition of language representation and control. *J. Neurolinguistics* **20**, 242–275 (2007).
20. Liu, L., He, Y. & Zheng, A. English and spanish bilinguals' language processing: An ALE-based meta-analysis of neuroimaging studies. in *Pattern recognition and computer vision* (eds. Liu, Q. et al.) 115–125 (Springer Nature Singapore, 2024).
21. Blanco-Elorrieta, E. & Pylkkänen, L. Bilingual language control in perception versus action: MEG reveals comprehension control mechanisms in anterior cingulate cortex and domain-general control of production in dorsolateral prefrontal cortex. *J. Neurosci.* **36**, 290–301 (2016).
22. Branzi, F. M., Rosa, P. A. D., Canini, M., Costa, A. & Abutalebi, J. Language control in bilinguals: Monitoring and response selection. *Cereb. Cortex* **26**, 2367–2380 (2016).
23. Walenski, M., Europa, E., Caplan, D. & Thompson, C. K. Neural networks for sentence comprehension and production: An ALE-based meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* **40**, 2275–2304 (2019).
24. Cox, R. W. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
25. Song, S. *et al.* Temporal similarity perfusion mapping: A standardized and model-free method for detecting perfusion deficits in stroke. *PLoS ONE* **12**, e0185552 (2017).

26. Olulade, O. A. *et al.* The neural basis of language development: Changes in lateralization over age. *Proc. Natl. Acad. Sci.* **117**, 23477–23483 (2020).
27. Ursache, A. & Noble, K. G. Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology* **53**, 71–82 (2016).
28. Hilger, K., Ekman, M., Fiebach, C. J. & Basten, U. Intelligence is associated with the modular structure of intrinsic brain networks. *Sci. Rep.* **7**, 16088 (2017).
29. Arredondo, M. M., Hu, X.-S., Seifert, E., Satterfield, T. & Kovelman, I. Bilingual exposure enhances left IFG specialization for language in children. *Biling. Lang. Cogn.* **22**, 783–801 (2019).
30. Kovelman, I., Baker, S. & Petitto, L. Bilingual and monolingual brains compared: An fMRI investigation of syntactic processing and a possible "neural signature" of bilingualism. *J. Cogn. Neurosci.* **20**, 153–169 (2008).
31. Roncaglia-Denissen, M. P. & Kotz, S. A. What does neuroimaging tell us about morphosyntactic processing in the brain of second language learners?. *Biling. Lang. Cogn.* **19**, 665–673 (2016).
32. Weber, K., Luther, L., Indefrey, P. & Hagoort, P. Overlap and differences in brain networks underlying the processing of complex sentence structures in second language users compared with native speakers. *Brain Connect.* **6**, 345–355 (2016).
33. Liu, X., Qu, J., Li, H., Yang, R. & Mei, L. Similar activation patterns in the bilateral dorsal inferior frontal gyrus for monolingual and bilingual contexts in second language production. *Neuropsychologia* **156**, 107857 (2021).
34. Połczyńska, M. M. & Bookheimer, S. Y. General principles governing the amount of neuroanatomical overlap between languages in bilinguals. *Neurosci. Biobehav. Rev.* **130**, 1–14 (2021).
35. Aben, B., Calderon, C. B., den Bussche, E. V. & Verguts, T. Cognitive effort modulates connectivity between dorsal anterior cingulate cortex and task-relevant cortical areas. *J. Neurosci.* **40**, 3838–3848 (2020).
36. Leon Guerrero, S., Whitford, V., Mesite, L. & Luk, G. Text complexity modulates cross-linguistic sentence integration in L2 reading. *Front. Commun.* **6**, 115 (2021).
37. August, D., Artzi, L. & Barr, C. Helping ELLs meet standards in English language arts and science: An intervention focused on academic vocabulary. *Read. Writ. Q.* **32**, 373–396 (2016).
38. Van Orman, D. S. J., Ardasheva, Y., Carbonneau, K. J. & Firestone, J. B. Examining the impacts of extended vocabulary instruction in mixed-English-proficiency science classrooms. *J. Educ. Res.* **114**, 74–88 (2021).
39. Luk, G. & Bialystok, E. Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *J. Cogn. Psychol.* **25**, 605–621 (2013).
40. Peirce, J. W. PsychoPy—Psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).
41. Kaufman, A. S. & Kaufman, N. L. *Kaufman Brief Intelligence Test* 2nd edn. (American Guidance Service, 2004).
42. Woodcock, R. W., Muñoz-Sandoval, A. F., Ruef, M. L. & Alvarado, C. G. *Woodcock-Muñoz Language Survey–Revised*. (Riverside, 2005).
43. Schrank, F., Wendling, B. & Alvarado, C. Woodcock-Muñoz Language Survey–Revised Normative Update. *Roll. Meadows IL Riverside* (2010).
44. Woodcock, R. W., Muñoz-Sandoval, A. F., Ruef, M. L. & Alvarado, C. G. *Bateria III Woodcock-Munoz: Pruebas de Habilidades Cognitivas* (Riverside Publishing Company, 2005).
45. Alvarado, R., Ruef, M. L. & Schrank, F. A. *Woodcock-Munoz Language Survey-Revised* (Riverside Publishing Company, 2005).
46. Zeileis, A., Wiel, M. A., Hornik, K. & Hothorn, T. Implementing a class of permutation tests: The coin package. *J. Stat. Softw.* **28**, 1–23 (2008).
47. Canty, A. J. Resampling methods in R: The boot package. *Newsl. R Proj.* **2**, 3 (2002).
48. Feinberg, D. A. *et al.* Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS ONE* **5**, e15710 (2010).
49. Moeller, S. *et al.* Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* **63**, 1144–1153 (2010).
50. Xu, J. *et al.* Evaluation of slice accelerations using multiband echo planar imaging at 3T. *NeuroImage* **83**, 991–1001 (2013).
51. Esteban, O. *et al.* fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
52. Gorgolewski, K. *et al.* Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Front. Neuroinformatics* **5**, 13 (2011).
53. Tustison, N. J. *et al.* N4ITK: Improved N3 bias correction. *IEEE Trans. Med. Imaging* **29**, 1310–1320 (2010).
54. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).
55. Fonov, V., Evans, A., McKinstry, R., Almli, C. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, S102 (2009).
56. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* **20**, 45–57 (2001).
57. Huntenburg, J. M. Evaluating nonlinear coregistration of BOLD EPI and T1w images. (Freie Universität Berlin, 2014).
58. Wang, S. *et al.* Evaluation of field map and nonlinear registration methods for correction of susceptibility artifacts in diffusion MRI. *Front. Neuroinformatics* **11**, 17 (2017).
59. Treiber, J. M. *et al.* Characterization and correction of geometric distortions in 814 diffusion weighted images. *PLOS ONE* **11**, e0152472 (2016).
60. Pruim, R. H. R. *et al.* ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage* **112**, 267–277 (2015).
61. Smith, S. M. & Brady, J. M. SUSAN—A new approach to low level image processing. *Int. J. Comput. Vis.* **23**, 45–78 (1997).
62. Ciric, R. *et al.* PennBBL/xcpEngine. Penn BBL (2019).
63. Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C. & Taylor, P. A. FMRI clustering in AFNI: False-positive rates Redux. *Brain Connect.* **7**, 152–171 (2017).
64. Cox, R. W. Equitable thresholding and clustering: A novel method for functional magnetic resonance imaging clustering in AFNI. *Brain Connect.* **9**, 529–538 (2019).
65. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Acknowledgements

The first author gratefully acknowledges support from the National Science Foundation Social and Behavioral Directorate (NSF SMA-2204272) during the writing of this manuscript. This research was carried out in part at the Harvard Center for Brain Science. This work involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program; specifically, grant number S10OD020039. This work was also partially funded using the Harvard University Mind Brain Behavior Graduate Student Award awarded to the second author. We acknowledge the financial support provided by the Harvard Graduate School of Education and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2020-05052) to the third author. The authors would like to thank Peter Shapiro at the Harvard Graduate School of Education for his assistance in producing the videos used in this study and to all the study team members for their assistance with data collection.

Author contributions

All authors contributed to the conception and design of the study, to acquiring and interpreting the data, and to drafting and editing the manuscript. S.L.G. and L.M. analyzed the data. S.L.G. wrote the first draft. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-69575-1.

Correspondence and requests for materials should be addressed to G.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024