TITLE: Language learning experience and mastering the challenges of perceiving speech in noise

AUTHORS & AFFILIATIONS: Shanna Kousaie[1,2], Shari Baum[2,3], Natalie A. Phillips[2,4,5], Vincent Gracco[2,3,6], Debra Titone[2,7], Jen-Kai Chen[1,2], Xiaoqian J. Chai[8], and Denise Klein[1,2]

[1]Cognitive Neuroscience Unit, Montreal Neurological Institute, McGill University, Montreal, QC, H3A 2B4, Canada
[2]Centre for Research on Brain, Language and Music, McGill University, Montreal, QC, H3G 2A8, Canada
[3]School of Communication Sciences and Disorders, Faculty of Medicine, McGill University, Montreal, QC, H3A 1G1, Canada
[4]Department of Psychology/Centre for Research in Human Development, Concordia University, Montreal, QC, H4B 1R6, Canada
[5]Bloomfield Centre for Research in Aging, Lady Davis Institute for Medical Research and Jewish General Hospital/McGill University Memory Clinic, Jewish General Hospital, Montreal, QC, H3T 1E2, Canada
[6]Haskins Laboratories, New Haven, CT, 06511, USA
[7]Department of Psychology, McGill University Montreal, QC, H3A 1G1, Canada
[8]Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, QC, H3A 2B4, Canada

CORRESPONDING AUTHORS:
Shanna Kousaie, Ph.D.
Montreal Neurological Institute, McGill University
3801 University Street
Montreal, Quebec
H3A 2B4
(514) 398-3372
shanna.kousaie@mail.mcgill.ca

Denise Klein, Ph.D.
Montreal Neurological Institute, McGill University
3801 University Street
Montreal, Quebec
H3A 2B4
(514) 398-3134
denise.klein@mcgill.ca

Abstract

Given the ubiquity of noisy environments and increasing globalization, the necessity to perceive speech in noise in a non-native language is common and necessary for successful communication. In the current investigation, bilingual individuals who learned their non-native language at different ages underwent magnetic resonance imaging while listening to sentences in both of their languages, in quiet and in noise. Sentence context was varied such that the final word could be of high or low predictability. Results show that early non-native language learning is associated with superior ability to benefit from contextual information behaviourally, and a pattern of neural recruitment in the left inferior frontal gyrus that suggests easier processing when perceiving non-native speech in noise. These findings have implications for our understanding of speech processing in non-optimal listening conditions and shed light on how individuals navigate every day complex communicative environments, in a native and non-native language.

KEYWORDS: age of acquisition; bilingualism; functional magnetic resonance imaging (fMRI); language; speech perception in noise

1. Introduction

Perceiving speech in noise is difficult (e.g., Cherry, 1953). Successful understanding of speech in noise is necessary for effective communication and represents a constant challenge to individuals given that most natural environments are noisy. Under certain circumstances, for instance when trying to process information in a non-native language (Lucks Mendel & Widner, 2016; Mayo, Florentine, & Buus, 1997; Shi, 2010; Tabri, Chacra, & Pring, 2011), perceiving speech in noise can be rendered more difficult. Increasing our understanding of the neural underpinnings of perceiving speech in noise in native and non-native languages is thus of fundamental importance given the increasing number of individuals who speak more than one language worldwide who function in their non-native language on a daily basis at work, school, or in social situations (e.g., Fabbro, 1999; Grosjean, 1989, 2008).

Some models of speech comprehension have been proposed to explain how the two languages of bilinguals interact to influence language processing and speech comprehension. For example, the Bilingual Language Interaction Network for Comprehension of Speech model (BLINCS; Shook & Marian, 2013) accounts for the fact that knowing more than one language can have an impact on the neurological or cognitive mechanisms involved in speech comprehension as a result of the simultaneous activation of a bilingual's two languages. The BLINCS model, focusses on the non-selectivity of language processing in bilinguals, and assumes shared phonological representations, integration across languages at the lexical level, and a single semantic system with shared conceptual representations. However, current models such as BLINCS do not comprehensively address bilingual speech perception in suboptimal listening conditions and the specific behavioural and/or neural consequences associated with functioning in a non-native language in terms of processing speech in noisy environments remain poorly understood.

Other models have focussed on the processing of speech in suboptimal listening conditions. For example, the Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016) postulates that in the context of understanding speech there is an interaction between effort, demand and motivation, and when the ratio between effort and reward becomes unacceptable (e.g., the effort required to comprehend speech outweighs reward), there may be a loss in motivation or "quitting" of task performance. However, this model does not directly address the processing of speech in a non-native language.

It has been suggested that functioning in a non-native language depletes cognitive resources more quickly than functioning in a native language, resulting in decreases in performance and productivity (Volk, Köhler, & Pudelko, 2014). Factors that could contribute to this include poorer phonetic discrimination (e.g., Bradlow & Alexander, 2007) and less automatic language processing (e.g., McLaughlin, Rossman, & McLeod, 1983; Mercier, Pivneva, & Titone, 2014; Segalowitz & Segalowitz, 1993) in a non-native language, as well as greater difficulty using contextual information to support speech perception.

In a native language, behavioural studies have shown that context facilitates the processing of speech in noise (e.g., Boothroyd & Nittrouer, 1988; Duffy & Giolas, 1974; Miller, 1962). For example, Miller, Heise, and Lichten (1951) found that words could be more accurately identified at lower signal-to-noise ratios (SNR; i.e., in more noise) when they were presented in a sentence (i.e., in context) compared to when they were presented in isolation. However, in terms of bilingualism several studies have found that bilingual participants perform more poorly in terms of perceiving speech in noise in their non-native language compared to monolinguals who are processing the information in their native language (Lucks Mendel & Widner, 2016; Mayo et al., 1997; Shi, 2010; Tabri et al., 2011). Furthermore, some have suggested that the age at which the non-native language is acquired could moderate the processing of speech in noise in the non-native language, with earlier acquisition being associated with native-like performance (Reetzke, Lam, Xie, Sheng, & Chandrasekaran, 2016; Tabri et al., 2011). Others have found that individuals who have attained a high level of proficiency in their non-native language perform similarly to their monolingual counterparts for stimuli that are not presented in context, but do not benefit from contextual information to the same extent as monolinguals (Skoe & Karayanidi, 2018). A behavioural study that directly compared the identification of words in noise in a native and non-native language within individuals found that participants performed better in their native compared to their non-native language and only benefitted from semantic information in their native language (Golestani, Rosen, & Scott, 2009). Conversely, in the non-native language, the presence of supportive semantic context resulted in unexpectedly poorer performance compared to when there was no context, suggesting that semantic information interfered with speech processing in the non-native language.

The neural underpinnings of processing speech in noise in a native language have been investigated using task-based and event-related functional magnetic resonance imaging (fMRI) under suboptimal listening conditions (e.g., spectral degradation, noise-vocoding). These studies have implicated the left and right temporal lobes, the left inferior frontal gyrus (LIFG), and the left angular gyrus (Clos et al., 2014; Davis, Ford, Kherif, & Johnsrude, 2011; Golestani, Hervais-Adelman, Obleser, & Scott, 2013; Obleser & Kotz, 2010; Obleser, Wise, Dresner, & Scott, 2007). Successful but effortful speech processing has been examined by presenting participants with sentences that varied in terms of semantic predictability and acoustic signal characteristics (Obleser & Kotz, 2010; Obleser et al., 2007). In these studies, when there was little semantic information, causing processing effort to be high, the LIFG was more active. Moreover, activity in inferior parietal cortex (supramarginal and angular gyri) was shown to be related to the availability of semantic information and the quality of the acoustic signal, suggesting an important role for these brain regions in facilitating the use of contextual information when the speech signal is compromised (Obleser & Kotz, 2010).

To date, only one fMRI study has compared the neural responses associated with processing speech in noise in a native and a non-native language. Hervais-Adelman, Pefkou, and Golestani (2014) observed that increased activity in the left angular gyrus for semantically-related compared to unrelated targets was restricted to the native language as compared to the non-native language. The left angular gyrus has previously been shown to be involved in different types of semantic tasks (e.g., Diaz & McCarthy, 2007; Noonan, Jefferies, Visser, & Lambon Ralph, 2013; Seghier, Fagan, & Price, 2010), and the finding of Hervais-Adelman et al. (2014) suggests that this involvement may be limited to a native language. Hervais-Adelman et al. (2014) also found that in the native language, there was greater recruitment for related compared to unrelated stimuli at higher relative to lower SNRs in the right cerebellum, the left planum temporale, and the left postcentral gyrus. Of note, the opposite pattern was observed in the non-native language (i.e., greater recruitment of the right cerebellum, the left planum temporale, and the left postcentral gyrus for related than unrelated trials at lower relative to higher SNRs).

In sum, while a small number of studies have examined the neural consequences of processing speech in noise in a non-native language; none have examined how the timing of language learning impacts these processes. Thus, the goal of the current investigation was to determine whether non-native speech processing differs as a result of language experience. Based on previous research demonstrating different developmental trajectories associated with the timing of non-native language learning in some brain regions implicated in language processing (e.g., Berken, Chai, Chen, Gracco, & Klein, 2016; Kaiser et al., 2015; Klein, Mok, Chen, & Watkins, 2014; Wei et al., 2015), we hypothesized that the age of non-native language learning would be associated with different patterns of neural recruitment during speech processing in suboptimal listening conditions in a non-native language. We also expected to observe behavioural differences as a function of the age of non-native language learning, in line with previous research (Reetzke et al., 2016; Tabri et al., 2011), with earlier learning being associated with better performance and later learning being associated with poorer use of contextual information.

In the current study, we had the opportunity to capitalise on the unique language learning environment of Quebec, where individuals learn their native language and non-native, second language at different points in development, and combined this with task-based fMRI to examine behavioural and brain differences associated with speech processing under challenging conditions. The strength of the study is that we compared individuals who have two native languages (i.e., individuals who had learned two languages from birth) to individuals who have high proficiency, but who varied with respect to the age at which they learned their non-native language. We manipulated the difficulty of speech processing by varying the degree of contextual constraint preceding target stimuli, thus making targets predictable or not, and by varying the language of the task such that participants processed speech in both their native and non-native languages. These manipulations allowed us to ascertain how individuals deal with speech in suboptimal listening conditions in a native and non-native language and how individual differences in language experience contribute to speech processing in a non-native language.

2. Materials and Methods

2.1 Participants

Thirty right-handed subjects who spoke only English and French were scanned. One group learned these two languages as native languages from birth (simultaneous group; n=10), one group learned the second language relatively early in life but after the first language (early group; n=12; mean age of acquisition=4.7 years, range=3-5 years), and a third group learned the second language after reaching school-age (late group; n=8; mean age of acquisition=6.9 years, range=6-9 years). For comparison with participants who learned the two languages sequentially, simultaneous bilinguals were asked to choose a nominal "native" language based on which language they were more comfortable in or used more frequently. At the time of testing, all participants were highly proficient in both English and French, did not differ in terms of objective measures of second language proficiency (i.e., letter and category fluency; proficiency tests are described in Supplementary Material) or relative usage of each language in terms of self-reported proportion of total language usage or frequency of usage[1]; however, all three groups differed from each other in terms of age of non-native language acquisition. The three groups were also matched in terms of chronological age, formal education, and general

[1] One-way ANOVAs were conducted on non-native language scores and self-reported language usage (in terms of proportion of total language usage and frequency of usage) and self-reported measures of codeswitching to ensure that there were no group differences. There were no significant effects of group for any of the objective language proficiency measures, or for the self-reported language usage measures (all *p*s > .16).

intelligence. The demographic information is reported in Table 1. All participants had pure-tone hearing thresholds within the normal range, self-reported good health, did not have knowledge of any languages other than French and English, had no history of traumatic brain injury or neurological disorder, or any other medical conditions or medications known to affect cognitive functioning, and did not have any conditions incompatible with MRI (e.g., metal implants, braces, electronically, magnetically, or mechanically activated devices such as cochlear implants, or claustrophobia). Given that musical training has been associated with differences in brain organization (Gaser & Schlaug, 2003) and improved speech perception in noise (Du & Zatorre, 2017), participants were all non-musicians.

INSERT TABLE 1

2.2 Stimuli and Materials

In addition to the battery of behavioural tests used to determine that the groups were matched across language, working memory and general intelligence (see Supplementary Material; data reported in Table 1), participants completed a speech perception in noise task in the MRI scanner. This experimental task was based on the revised Speech Perception in Noise task (SPIN-R; Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984). Sentences were presented to participants through MRI-compatible Sensimetrics S14 insert earphones (Sensimetrics Corporation) with Comply foam canal tips (Hearing Components, Oakdale, MN) while they were in the MRI scanner and participants were instructed to repeat out loud the final word of the sentence following a prompt during a quiet interscan interval. Sentences were either presented in quiet or in noise (i.e., multi-talker babble) with a signal-to-noise ratio of -6 decibels and the final word was either of high or low predictability (i.e., the preceding sentence provided high or low context for the final word). The multi-talker babble was adapted from the original eight-talker babble from the SPIN-R (Bilger et al., 1984) by overlaying the original eight-talker babble track three times slightly shifted in time, thus creating a multi-talker babble that was less variable in terms of intensity fluctuations (Winneke & Phillips, 2011). Equivalent French and English language sentences were created with 27 sentences in each of 8 experimental conditions (i.e., 4 conditions in each language; high predictability in quiet, high predictability in noise, low predictability in quiet, and low predictability in noise). Each sentence terminal word was heard in all possible conditions within each language; however, two equivalent lists were created such that terminal words repeated only twice in each list, once in a high predictability sentence and once in a low predictability sentence, and once in noise, once in quiet (e.g., the terminal word *spoon* was heard in high predictability quiet and low predictability noise in list 1, and in low predictability quiet and high predictability noise in list 2). Each participant heard only one list. Sentence length was matched across high and low predictability conditions, and the final word of each sentence (the target word) was always monosyllabic and was matched across languages in terms of spoken frequency, phonological neighbourhood density, imageability, and familiarity using the MRC Psycholinguistic Database (Coltheart, 1981), Lexique 3 (New, 2006; New, Pallier, Ferrand, & Matos, 2001), and the Corpus of Contemporary American English (Davies, 2008-). See Table 2 for sample stimuli.

INSERT TABLE 2

Sentences were digitally recorded using an Olympus LS-11 Linear PCM recorder at a rate of 44.1 kHz with 16-bit resolution. The recordings in English and French were made by a female simultaneous bilingual speaker who was judged native-like in her speech by native speakers in each language. Stimuli were edited offline using Audacity recording and editing software version 2.0.5 (Audacity Team, 2014); this included normalization and compression to ensure that the peak amplitude and dynamic range of the stimuli were similar across all recordings. Within each list, sentences were blocked by language and by noise, low and high predictability sentences were intermixed within each block, and the blocks were separated by four baseline trials during which participants fixated on a cross and heard either silence or multi-talker babble. The speech perception task was presented using E-Prime 2.0 (SP1) presentation software (Psychology Software Tools, Pittsburgh, PA, USA). Responses were recorded using an Olympus LS-11 Linear PCM recorder and scored for accuracy offline.

2.3 Experimental and Scanning Procedure

Participants completed a behavioural testing session lasting 90 minutes, followed by a scanning session (2.5 hours). To ensure that participants understood the task, prior to scanning they were trained on the speech perception task outside of the scanner using sentences not included in the experimental task. Participants were instructed to listen to the sentence and wait for the beep indicating that they should repeat out loud the final word of the sentence. Each trial lasted 7000 ms and started with sentence presentation (3000 ms), followed by scan acquisition (2260 ms) and the response interval (1740 ms). The length of the sentences varied between 1405 ms and 2812 ms and stimuli were presented such that the end of sentence was immediately followed by scan acquisition, which was followed by the response beep indicating that the participant should respond. This sparse-sampling method and timing was implemented to ensure that sentences were presented during the quiet interval between scan acquisitions. Ethics approval for this study was obtained from the Research Ethics Board at the Montreal Neurological Institute (MNI), McGill University, and participants gave their written consent.

Imaging was performed at the MNI on a 3T TrioTim Siemens scanner using a 32-channel head coil. The speech perception task was completed in two runs lasting 14:44 minutes each. Functional images were acquired with a T_2*-weighted gradient echo-planar imaging sequence (EPI) in 38 3.5mm thick transverse slices covering the entire brain (TR=7000 ms, TA=2260 ms, TE=30 ms, FoV=224 mm, flip angle=90 degrees, interleaved excitation); a total of 125 volumes were obtained in each run. High-resolution T_1-weighted images were obtained from a 3D magnetization prepared rapid acquisition gradient echo (MP-RAGE) sequence (slice thickness=1 mm, TR=2300 ms, TE=2.98 ms, matrix size=256 x 256, FoV=256 mm, flip angle=9°, interleaved excitation) for each participant and used as an anatomical reference. The speech perception task was part of a larger study in which participants also completed two additional unrelated functional tasks, resting-state fMRI, and diffusion-weighted imaging.

3. Analysis and Results

3.1 Behavioural results

Behavioural data were analyzed with an analysis of variance (ANOVA) with the between subjects factor Group (simultaneous, early, late), and within subjects factors Language (native, non-native)[2], Predictability (high, low), and Noise (noise, quiet). Error rates were larger in participants' non-native compared to their native language (main effect of Language;

---

[2] Given that all of our participants were highly proficient in their non-native language based on self-report and objective measures, and that in the individuals who learned both languages simultaneously the native language was nominal, we created language groups based on native language, irrespective of whether the native language was English or French. As can be seen in Table 1, seven participants reported French as their native language.

$F(1,27)=22.8$, $MSE=0.12$, $\eta^2_p=.46$, $p<.01$), in response to low than high predictability stimuli (main effect of Predictability; $F(1,27)=120.7$, $MSE=0.01$, $\eta^2_p=.82$, $p<.01$), and in noise compared to quiet (main effect of Noise; $F(1,27)=302.4$, $MSE=0.03$, $\eta^2_p=.92$, $p<.01$). There was also an interaction between Language, Predictability, and Noise ($F(1,27)=13.2$, $MSE=0.01$, $\eta^2_p=.33$, $p<.01$), showing that error rates were larger for low than high predictability stimuli in noise only, and that the effect of Language (i.e., higher error rates in the non-native compared to the native language) was larger for high predictability stimuli. These data are shown in Figure 1.

In the quiet condition participants were performing almost at ceiling (i.e., all error rates less that 2.7%), indicating that in favourable listening conditions participants were able to accurately perceive both languages, thus, we examined each noise condition separately. Specifically, we conducted an ANOVA with the between subjects factor Group (simultaneous, early, late), and within subjects factors Language (native, non-native) and Predictability (high, low) for noise and for quiet separately. In quiet, there was a main effect of Language ($F(1,27)=5.6$, $MSE=0.001$, $\eta^2_p=.17$, $p=.03$) showing that responses were less accurate in the non-native compared to the native language. There were no other significant effects. In noise, there were main effects of both Language ($F(1,27)=23.1$, $MSE=0.02$, $\eta^2_p=.46$, $p<.001$) and Predictability ($F(1,27)=102.7$, $MSE=0.01$, $\eta^2_p=.79$, $p<.001$), as well as a significant interaction between Language and Predictability ($F(1,27)=12.1$, $MSE=0.02$, $\eta^2_p=.31$, $p<.01$). These findings indicated that responses were less accurate in the non-native compared to the native language, for low compared to high predictability stimuli, and that the effect of Language was larger for high predictability stimuli than for low predictability stimuli. Given that we hypothesized that individuals who learned their non-native language late would show smaller predictability effects than individuals who learned their non-native language earlier, we examined the simple effect of

Predictability in the Language Group x Language x Predictability interaction (p=.12). This analysis revealed that there was a significant effect of Predictability (less accurate responses for low compared to high predictability stimuli) for all groups in both languages (all ps≤.01) with the exception of the late group, where there was no predictability effect in the non-native language (p=.22).[3]

INSERT FIGURE 1

3.2 Task-based functional MRI results

Functional imaging data were preprocessed and analyzed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). Preprocessing followed standard steps, including realignment and unwarping, slice time correction, segmentation, normalization in MNI space and smoothing with a 6mm full width at half maximum (FWHM) Gaussian kernel. Artifact/outlier scans were defined as images in which average intensity deviated more than 3 standard deviations from the mean intensity in the session, or composite head movement exceeded 1.5mm from the previous image. A total of 5.9% of trials were identified as artifacts/outliers and were removed from the analysis using ART (Artifact Detection Tools; http://nitrc.org/projects/artifact_detect/).

[3] It is noteworthy that in Figure 1 it appears as though the individuals who learned their non-native language late show a larger predictability effect in noise in the native language. In order to address this, we ran a Language Group x Predictability ANOVA for the noise condition in the native language and found a main effect of Predictability (more accurate responses for high compared to low predictability) and no other significant effects.

As an initial step, we examined the brain regions recruited for processing speech in noise compared to a silent baseline for each condition. Not unexpectedly, similar networks of regions were recruited for all conditions involving processing speech in noise (see Figure 2; see also Supplementary Figure 1 which shows the brain regions that were recruited for the 8 experimental conditions for each group and language separately). To address our specific hypotheses, we examined whether there were brain regions that were differentially recruited in noise compared to quiet, and in low compared to high predictability stimuli in noise in the native and non-native language, and whether effects were related to the age at which the non-native language was learned. Given the number of experimental conditions, and based on our behavioural findings, we created first level contrasts for each participant to compare noise and quiet conditions and low and high predictability stimuli in noise. At the second level, we conducted Group x Language ANOVAs to examine main effects and interactions between Group and Language for the noise vs. quiet and high predictability vs. low predictability contrasts individually. Our primary hypothesis was that neural recruitment would differ for listening to non-native speech in suboptimal listening conditions as a function of the timing of non-native language learning. We addressed this hypothesis in a Group x Language ANOVA with the first-level contrasts comparing high and low predictability stimuli in noise. We also examined the average across all 30 participants separately for each language to determine if there were brain regions showing significantly greater activity for low than high predictability stimuli. All reported effects were significant at an uncorrected $p<.001$, as recommended by Woo, Krishnan, and Wager (2014), with a cluster extent threshold of 100 voxels, and survived cluster-level correction at $p(\text{FDR})<.05$. The cluster extent threshold that we used was based on previously published studies examining speech perception in noise using fMRI methodology (e.g., Hervais-Adelman et al., 2014; Obleser & Kotz, 2010; Obleser et al., 2007).

INSERT FIGURE 2

3.2.1 Noise vs Quiet. The Group x Language ANOVA revealed no significant effects.

3.2.2 Low vs high predictability. The Group x Language ANOVA revealed no main effects; however, there was a significant interaction between Group and Language in a region of the anterior LIFG corresponding to Brodmann area 47 (see Figure 3, panel A). We decomposed the interaction by looking at the effect of Group in each language separately in a one-way ANOVA and found a significant effect of Group in the non-native language only. Specifically, the relative recruitment of the LIFG differed as a function of when the non-native language was learned. In order to determine the source of this group difference, we used the *rex* tool in MATLAB to extract blood-oxygen-level dependent (BOLD) activity from the specific LIFG cluster in which a group difference in activation for the low predictability minus high predictability contrast was observed for each participant and these data were subsequently subjected to a one-way ANOVA with Bonferroni corrected posthoc comparisons.

This analysis showed that all three groups differed from each other ($F(2,27)=24.2$, $MSE=1.9$, $p<.01$). As a final step, a one-sample t-test for each group was conducted to determine if the observed difference in BOLD signal for low compared to high predictability stimuli differed from zero. The t-test revealed that in individuals who had learned both of their languages from birth, the LIFG was significantly more activated in response to low than high predictability stimuli ($t(9)=4.9$, $p<.01$), whereas in individuals who learned their second language after school-age, an opposite pattern emerged, with greater LIFG activity in response to high than low predictability stimuli ($t(9) = -5.6$, $p < .01$). Finally, in the group who had learned their second language early, there was no difference in the recruitment of the LIFG for the two conditions (i.e., the difference in BOLD signal for the low compared to high predictability conditions did not differ from 0; $t(11) = -1.6$, $p = .13$). See Figure 3, panel B.

While there were no regions that showed an effect of Group in the native language, we did find a cluster in the LIFG that was active irrespective of group. Specifically, when we examined the average activity across all participants a cluster in the LIFG was found to be recruited to a greater extent for low than high predictability stimuli in the native language, see Figure 3, panel C. Results of the fMRI analyses are reported in Table 3.

INSERT FIGURE 3

INSERT TABLE 3
----

4. Discussion and Conclusions

Most listening environments are noisy (e.g., ambient background noise in classrooms, open-concept offices, public places) and a large proportion of individuals must function in these environments on a daily basis, many doing so in their non-native language. In natural environments, individuals are able to use contextual information to extract the speech signal from noise in their native language. However, in a non-native language and when there is no contextual information, processing speech in noise becomes more challenging – situations in which individuals commonly find themselves. In the current investigation, we report converging evidence from behaviour and fMRI that there are both behavioural and neural consequences related to language learning experiences, namely that the timing of non-native language learning has an impact on the processing of speech in a non-native language in sub-optimal listening conditions, with later learning being associated with greater difficulty and differential recruitment of neural resources.

The behavioural results show that performance was better in quiet than in noise, for predictable than non-predictable stimuli, and in a native compared to a non-native language. Improved performance for predictable compared to unpredictable stimuli was limited to situations in which the stimuli were presented in noise; that is, when speech perception became more difficult, the benefit of supporting contextual information emerged. This was true in both the native and the non-native language, although the benefit was larger in the native language, demonstrating that participants benefitted from contextual information to a greater extent in their native than their non-native language. This is consistent with the recent findings of Skoe and Karayanidi (2018), who showed that non-native language participants did not benefit from contextual information to the same extent as monolinguals did. Although Skoe and Karayanidi compared different groups of participants, our findings are similar in that we show a smaller benefit of context in a non-native language as compared to a native language.

Of particular interest is the finding that the benefit of contextual information in the non-native language depended on when the non-native language was learned. That is, individuals who learned their non-native language after the age of six years did not appear to benefit from contextual information when processing speech in noise in their non-native language. Thus, an individual's ability to use contextual information to process challenging speech is likely associated with the timing of language learning. This is further supported by our task-based fMRI data, where we see differential recruitment of the anterior LIFG for challenging speech processing in a non-native language only. In the native language, and irrespective of when the non-native language was learned, the LIFG was recruited to a greater extent in response to low compared to high predictability stimuli, which is consistent with findings showing that the LIFG is important for processing speech when processing is more effortful (Obleser & Kotz, 2010). However, in the non-native language, the LIFG was activated in distinct ways depending on the timing of second language learning, with greater recruitment of the LIFG for low (more challenging) compared to high predictability stimuli in individuals who learned their two languages from birth, and an opposite pattern in individuals who learned their non-native language after 6 years of age. Given that previous research has suggested that the LIFG is important for effortful speech processing (Obleser & Kotz, 2010), as well as for processing semantic information (e.g., Booth et al., 2002; Price, 2010), degraded speech (Davis & Johnsrude, 2003), and when the availability of bottom-up sensory information decreases (Shahin, Bishop, & Miller, 2009), our fMRI findings further suggest a role for this region for processing of speech in noise in a non-native language, which is more difficult when the non-native language is learned late. The pattern of LIFG activation, taken together with the behavioural findings, show that these individuals failed to show a processing benefit in the presence of contextual information. The results suggest that in late language learners as compared to their counterparts who learned their non-native language earlier, processing speech in noise in the non-native language may exhaust available resources at lower task demands (i.e., for high predictability stimuli). In this situation, the LIFG is recruited for high predictability stimuli, leaving these resources unavailable for the more difficult low predictability stimuli, and resulting in similarly poor performance for both high and low predictability stimuli.

Alternatively, it is possible that learning a non-native language following mastery of a native language results in less efficient auditory processing in the non-native language, making it more difficult to extract the speech signal from noise. This type of pattern has been reported by Obleser and Kotz (2010), who observed that the increased activity in the LIFG for low- compared to high-cloze sentences was greater as intelligibility of the sentences increased. This alternative interpretation is unlikely in the present study; the absence of any group differences in activity in auditory cortex in response to speech presented in noise in the non-native language suggests similar processing of the auditory signal across groups.

Another possibility is that extracting the speech signal from noise in a non-native language when there was little contextual information was so difficult for the individuals who had learned their non-native language after 6 years of age, that the effort invested in comprehending the sentences decreased, resulting in poor performance and less activation of the LIFG. This interpretation is consistent with the FUEL model (Pichora-Fuller et al., 2016), which proposes that when the effort required to comprehend speech outweighs reward, there is a loss in motivation and a decrease in task performance. Given that we did not collect subjective information regarding how effortful the task was for participants, it is difficult to assess the likelihood of this alternate explanation.

The LIFG has previously been implicated in semantic and sentence processing (Peelle, 2019). Furthermore, previous research using transcranial magnetic stimulation has suggested that the anterior LIFG, similar to the region in which we observe an effect of timing of non-native language learning on speech perception in noise, is implicated in sematic processing (Gough, Nobre, & Devlin, 2005). Consistent with this notion, the effect of predictability (i.e., the use of contextual information to facilitate the processing of speech in suboptimal listening conditions) that we observed in the anterior LIFG is likely related to the ability to extract semantic information from speech under difficult listening conditions.

Given that our sample was comprised of bilinguals who had learned their non-native language relatively early in life (i.e., before 10 years of age) with one third having learned their non-native language from birth we initially ran a group-based analysis comparing groups of bilinguals who had learned their two languages at different points in development (i.e., from birth, between 1 and 5 years of age, and after 6 years of age). However, in a supplemental analysis we examined the age of non-native language acquisition as a continuous variable by converting the age of acquisition to standardized z-scores. There was no relationship between age of acquisition and behavioural performance; however, earlier age of acquisition was related to greater recruitment of the anterior LIFG for low compared to high predictability stimuli in noise in the non-native language (see Supplementary Figure 2). This finding is consistent with the results from our group analysis, demonstrating that our findings hold even when our sample is not grouped by age of non-native language acquisition, and provides further support for the suggestion that the timing of non-native language learning has implications for the processing of non-native speech in noise.

Our data speak to the importance of understanding the processes important for perceiving speech in noise and provide some insight into how previous models of speech processing may be modified to provide a more refined representation of bilingual speech processing in suboptimal listening conditions. For example, the FUEL model (Pichora-Fuller et al., 2016) was not originally proposed to address how effortful listening may differ in a non-native compared to a native language, or as a function of non-native language experience, but it could be modified by considering non-native language listening and individual differences in language experience in the "input-related demands" component of the model which includes elements that contribute to unfavourable listening conditions. Similarly, given the computational nature of the BLINCS model (Shook & Marian, 2013) it seems possible that noise that simulates the effort required in noisy listening environments could be introduced. As well, training the self-organizing maps that are integral to the model to account for individual differences in language experience (e.g., Miikkulainen & Kiran, 2009; Zhao & Li, 2010) might better represent the bilingual language system in terms of suboptimal speech processing in a non-native language.

As a final caveat, our design was constrained by the noisiness of the fMRI scanning environment, which made it necessary to include a delay between sentence presentation and participant responses to ensure that sentences were presented during the quiet interval between scan acquisitions. This may have negatively impacted performance by introducing additional working memory demands as a result of the need for participants to maintain the sentence final word in working memory for the duration of the delay. Given that each sentence was heard in all experimental conditions, any demands placed on working memory by scanning parameters did not vary systematically between conditions. As well, there were no differences between the groups in terms of working memory ability (see Table 1). However, it is unknown whether increases in task difficulty (i.e., processing speech in suboptimal listening conditions and in a non-native language) introduced additional working memory demands. Additionally, it is possible that the timing of non-native language learning moderated the interaction between task difficulty and working memory demands. For example, the ability to process speech in noise may differ as a function of task difficulty such that when the task is easy (i.e., in optimal listening conditions) and when there are few demands on cognitive resources there is no effect of the sentence presentation-response delay on performance. Whereas, when demands on cognitive resources increase with increasing task difficultly, demands on working memory may be imposed and may affect the impact of the presentation-response delay on performance. The timing of non-native language learning could further influence the extent to which task difficulty increases demands on working memory.

In conclusion, our findings are consistent with previous research showing poorer behavioural speech perception in noise in the non-native language of bilinguals compared to monolinguals processing in their native language (Lucks Mendel & Widner, 2016; Mayo et al., 1997; Shi, 2010; Tabri et al., 2011), as well as in the non-native compared to the native language of bilinguals (Golestani et al., 2009). Furthermore, our fMRI results are consistent with findings suggesting that there are different neural mechanisms implicated in the use of contextual/semantic information in the processing of speech in suboptimal listening conditions in a native and non-native language (Hervais-Adelman et al., 2014). Critically, there were no group differences in non-native language proficiency, thus, our findings demonstrate that the timing of language learning has implications for processing speech in noise in a non-native language irrespective of attained proficiency in the non-native language. Furthermore, previous research shows that earlier age of non-native language acquisition is associated with decreased cortical thickness in the LIFG (Klein et al., 2014), greater functional connectivity of the LIFG (Berken et al., 2016), and a more integrated functional language network (Thieba, Long, Dewey, & Lebel, 2019); our findings suggest that these changes in brain structure and in intrinsic functional organization may have consequences for the functions subserved by the LIFG, such as difficult speech processing in a non-native language.

The current findings are relevant to discussions about how individuals deal with non-optimal listening conditions and increase our understanding of how individuals with differing language backgrounds navigate the complex communicative environments that they are faced with in their everyday lives. Importantly, our findings shed light on the impact of the timing of non-native language learning on the early stages of brain development in relation to non-native language processing. In the future, the inclusion of larger sample sizes and more detailed information on language use patterns may shed additional light on the processing of speech under challenging conditions in a non-native language.

5. Acknowledgements

The authors would like to thank Megan Callahan for her help translating the French stimuli and the technicians at the Brain Imaging Centre of the Montreal Neurological Institute for their help with MRI data collection.

6. Funding

This work was supported by a Canadian Institutes of Health Research postdoctoral fellowship to SK, by an individual Discovery Grant from the Natural Sciences and Engineering Research Council of Canada to DK, by a Fonds de recherches du Québec – Société et culture (FRQSC), Soutien aux équipes de recherche grant to SB, VG, DT, DK, NP, by the Fonds de recherches du Québec – Société et culture and Nature et technologies (FRQSC/FRQNT) funded Centre for Research on Brain Language and Music, and by funds from the Blema and Arnold Steinberg Family Foundation.

Declarations of interest: none

5. References

Audacity Team. (2014). Audacity(R): Free Audio Editor and Recorder version. Version 2.0.0 retrieved April 20th 2014 from http://audacity.sourceforge.net/.

Berken, J. A., Chai, X., Chen, J. K., Gracco, V. L., & Klein, D. (2016). Effects of early and late bilingualism on resting-state functional connectivity. *The Journal of Neuroscience, 36*(4), 1165-1172. doi:10.1523/jneurosci.1960-15.2016

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27*(1), 32-48.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2002). Modality independence of word comprehension. *Human Brain Mapping, 16*(4), 251-261. doi:10.1002/hbm.10054

Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America, 84*(1), 101-114.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *J Acoust Soc Am, 121*(4), 2339-2349.

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *J Acoust Soc Am, 25*(5), 975-979. doi:10.1121/1.1907229

Clos, M., Langner, R., Meyer, M., Oechslin, M. S., Zilles, K., & Eickhoff, S. B. (2014). Effects of prior information on decoding degraded speech: an fMRI study. *Human Brain Mapping, 35*(1), 61-74. doi:10.1002/hbm.22151

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A, 33*(4), 497-505. doi:10.1080/14640748108400805

Davies, M. (2008-). The corpus of contemporary American English (COCA): 560 million words, 1990-present. https://corpus.byu.edu/coca/

Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience, 23*(12), 3914-3932. doi:10.1162/jocn_a_00084

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience, 23*(8), 3423-3431.

Diaz, M. T., & McCarthy, G. (2007). Unconscious word processing engages a distributed network of brain regions. *Journal of Cognitive Neuroscience, 19*(11), 1768-1775. doi:10.1162/jocn.2007.19.11.1768

Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences of the United States of America, 114*(51), 13579-13584. doi:10.1073/pnas.1712223114

Duffy, J. R., & Giolas, T. G. (1974). Sentence intelligibility as a function of key word selection. *Journal of Speech and Hearing Research, 17*(4), 631-637.

Fabbro, F. (1999). *The neurolinguistics of bilingualism. An introduction.* Hove, East Sussex, UK: The Psychology Press.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience, 23*(27), 9240-9245.

Golestani, N., Hervais-Adelman, A., Obleser, J., & Scott, S. K. (2013). Semantic versus perceptual interactions in neural processing of speech-in-noise. *NeuroImage, 79*, 52-61. doi:10.1016/j.neuroimage.2013.04.049

Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. *Bilingualism: Language and Cognition, 12*(3), 385-392. doi:10.1017/s1366728909990150

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *The Journal of Neuroscience, 25*(35), 8010. doi:10.1523/JNEUROSCI.2307-05.2005

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3-15.

Grosjean, F. (2008). *Studying Bilinguals.* Oxford: Oxford University Press.

Hervais-Adelman, A., Pefkou, M., & Golestani, N. (2014). Bilingual speech-in-noise: Neural bases of semantic context use in the native language. *Brain and Language, 132*, 1-6. doi:10.1016/j.bandl.2014.01.009

Kaiser, A., Eppenberger, L. S., Smieskova, R., Borgwardt, S., Kuenzli, E., Radue, E.-W., . . . Bendfeldt, K. (2015). Age of second language acquisition in multilinguals has an impact on gray matter volume in language-associated brain areas. *Frontiers in Psychology, 6*, 638. doi:10.3389/fpsyg.2015.00638

Klein, D., Mok, K., Chen, J.-K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: A cortical thickness study of bilingual and monolingual individuals. *Brain and Language, 131*, 20-24. doi:http://dx.doi.org/10.1016/j.bandl.2013.05.014

Lucks Mendel, L., & Widner, H. (2016). Speech perception in noise for bilingual listeners with normal hearing. *International Journal of Audiology, 55*(2), 126-134. doi:10.3109/14992027.2015.1061710

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing Research, 40*(3), 686-693.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*(2), 135-158. doi:10.1111/j.1467-1770.1983.tb00532.x

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition, 17*(1), 89-117. doi:10.1017/S1366728913000084

Miikkulainen, R., & Kiran, S. (2009). Modeling the bilingual lexicon of an individual subject. *Lecture Notes in Computer Science, 5629*, 191-199.

Miller, G. A. (1962). Decision units in the perception of speech. *IRE Transactions on Information Theory, IT-8*, 81-83.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41*(5), 329-335.

New, B. (2006). Lexique 3: Une nouvelle base de données lexicales. *Actes de la Conférence Traitement Automatique des Langues Naturelles (TALN 2006)*. Louvain, Belgique.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Anné Psychologique, 101*, 447-462.

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience, 25*(11), 1824-1850. doi:10.1162/jocn_a_00442

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex, 20*(3), 633-640. doi:10.1093/cercor/bhp128

Obleser, J., Wise, R. J., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience, 27*(9), 2283-2289. doi:10.1523/jneurosci.4663-06.2007

Peelle, J. E. (2019). The neural basis for auditory and audiovisual speech perception. In W. F. Katz & P. F. Assman (Eds.), *The Routledge Handbook of Phonetics*: Routledge.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . . Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear & Hearing, 37*(S1), 5s-27s. doi:10.1097/aud.0000000000000312

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*, 62-88. doi:10.1111/j.1749-6632.2010.05444.x

Reetzke, R., Lam, B. P., Xie, Z., Sheng, L., & Chandrasekaran, B. (2016). Effect of simultaneous bilingualism on speech intelligibility across different masker types, modalities, and signal-to-noise ratios in school-age children. *PLoS One, 11*(12), e0168048. doi:10.1371/journal.pone.0168048

Segalowitz, N., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14*, 369-385.

Seghier, M. L., Fagan, E., & Price, C. J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *Journal of Neuroscience, 30*(50), 16809-16817. doi:10.1523/jneurosci.3377-10.2010

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage, 44*(3), 1133-1143. doi:10.1016/j.neuroimage.2008.09.045

Shi, L. F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of english acquisition. *Journal of Speech, Language and Hearing Research, 53*(4), 821-835. doi:10.1044/1092-4388(2010/09-0081)

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism, 16*(2), 304-324. doi:10.1017/s1366728912000466

Skoe, E., & Karayanidi, K. (2018). Bilingualism and speech understanding in noise: Auditory and linguistic factors. *J Am Acad Audiol*.

Tabri, D., Chacra, K. M., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders, 46*(4), 411-422. doi:10.3109/13682822.2010.519372

Thieba, C., Long, X., Dewey, D., & Lebel, C. (2019). Young children in different linguistic environments: A multimodal neuroimaging study of the inferior frontal gyrus. *Brain and Cognition, 134*, 71-79. doi:10.1016/j.bandc.2018.05.009

Volk, S., Köhler, T., & Pudelko, M. (2014). Brain drain: The cognitive neuroscience of foreign language processing in multinational corporations. *Journal of International Business Studies, 45*(7), 862-885. doi:10.1057/jibs.2014.26

Wechsler, D. (2008). *WAIS-IV technical and interpretive manual*. San Antonio, TX: The Psychological Corporation.

Wei, M., Joshi, A. A., Zhang, M., Mei, L., Manis, F. R., He, Q., . . . Lu, Z.-L. (2015). How age of acquisition influences brain architecture in bilinguals. *Journal of Neurolinguistics, 36*, 35-55. doi:10.1016/j.jneuroling.2015.05.001

Winneke, A. H., & Phillips, N. A. (2011). Does audiovisual speech offer a fountain of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception. *Psychology and Aging, 26*(2), 427-438. doi:10.1037/a0021683

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage, 91*, 412-419. doi:https://doi.org/10.1016/j.neuroimage.2013.12.058

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism, 13*(5), 505-524. doi:10.1080/13670050.2010.488284

Supplementary Material

Assessment of language proficiency. In addition to self-reporting their proficiency in their native and non-native language, participants completed letter fluency, and category fluency tasks in both languages in order to objectively measure proficiency; scores are reported in Table 1.

In the letter fluency tasks participants were asked to produce as many words as they could think of that started with a specific letter of the alphabet, excluding proper nouns, numbers, or words differing only in terms of their suffix (e.g., love, lover, loving). They were given one minute per letter to produce as many words as possible for three letters in each language (*F*, *A*, and *S* in English; *P*, *F*, and *L* in French) and their score reflects the total number of words produced in each language.

The category fluency task was similar to the letter fluency tasks; however, participants were required to produce as many exemplars as possible from a specific category in one minute (*animals* in English; *fruits* in French). Each participant's score reflects the total number of correct exemplars produced.

Assessment of working memory and general intelligence. The Digit Span, Letter Number Sequencing, and Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale Fourth Edition (Wechsler, 2008) were used to assess working memory and general intelligence/reasoning ability; standardized scores are reported in Table 1. The assessment of working memory and general intelligence was completed in the participant's native language.

The Digit Span task is comprised of Forward Digit Span, Backward Digit Span, and Sequencing. For each variant of the task the participant is required to repeat a series of numbers either in the forward order (Forward Digit Span), the backward order (Backward Digit Span), or in sequential order form smallest to largest (Sequencing). There are two trials at each span length and the span length increases by one digit each time the participant obtains at least one correct trial at each span length. The task is terminated when the participant obtains two incorrect responses at any given span length.

In the Letter Number Sequencing task participants are read a list of letters and numbers and are required to repeat the list back with the numbers first in sequential order from smallest to largest, followed by the letters in alphabetical order. There are three trials at each span length and the span increases by one digit or letter each time the participant obtains at least one correct trial at each span length. The task is terminated when the participant obtains three incorrect responses at any given span length.

In Matrix Reasoning, participants are presented with a series of designs and are required to identify the patterns. There is a total of 26 designs that increase in complexity/difficulty.

Figure Captions

Figure 1. Panels A-D show error rates as a function of Language Group and Predictability in quiet and in noise for each language. Error rates were larger in noise (panels C and D) than in quiet (panels A and B) and for low than high predictability sentences. The effect of predictability (higher accuracy for high than low predictability stimuli) was significant for all groups in both the native and the non-native language in noise, with the exception of the group who had learned their non-native language after age 6 where there was no predictability effect in the non-native language (p=.22).

Figure 2. Brain regions recruited for each of the four experimental conditions collapsed across native and non-native language; high predictability in quiet (panel A), high predictability in noise (panel B), low predictability in quiet (panel C), and low predictability in noise (panel D). A similar network of regions is implicated in speech processing across all conditions. Images are based on a whole brain analysis comparing task conditions to a silent baseline and a threshold of $p(\text{FWE}) < .05$. Note that we collapsed across the two languages given that there were no effects of Language. Supplementary Figure 1 shows the brain regions recruited for the 8 experimental conditions for each group and language separately.

Figure 3. Results from a whole-brain Language x Group analysis of variance (ANOVA) comparing relative activation for low compared to high predictability sentences. The F-map in panel A shows the region in the left anterior inferior frontal gyrus (LIFG) corresponding to Brodmann area 47 where a significant interaction between Group and Language was observed. Follow-up whole-brain one-way ANOVAs for each language showed differential recruitment of

the LIFG in the non-native language as a function of when the non-native language was learned. The F-map in Panel B shows the LIFG region in which a systematic decrease in relative activation was observed in response to low compared to high predictability stimuli in noise. The bar graph at the bottom of panel B shows the systematic relative decrease in BOLD signal (low predictability minus high predictability) in this LIFG region for each group in their non-native language. Individuals who learned their two languages simultaneously showed greater activation for low than high predictability stimuli, and the opposite pattern was observed in individuals who learned their non-native language after 6 years of age. Of note is that the negative BOLD signal observed for the late group in the non-native language does not represent deactivation, but rather that there was greater activation for high compared to low predictability stimuli. There were no group differences in the native language; however, the t-map in Panel C shows a region in the LIFG that was recruited in response to low compared to high predictability stimuli in noise in the native language across participants, irrespective of when they learned their non-native language. Peak coordinates for these regions are reported in Table 3.

Supplementary Figure 1. Regions of activation for the 8 experimental conditions for each group and language separately. Results are from a whole brain analysis comparing the condition of interest to a quiet baseline and are thresholded using uncorrected $p < .001$ and an extent threshold of 100 voxels.

Supplementary Figure 2. A whole-brain regression analysis with standardized age of acquisition scores revealed a region in the anterior left inferior frontal gyrus that was recruited to a greater extent for high compared to low predictability stimuli in noise in the non-native language in individuals with earlier non-native language acquisition, consistent with the group analysis (uncorrected $p < .001$, $k > 100$, $p\ (FDR) < .05$). The scatterplot shows the relationship between the difference in BOLD signal for low compared to high predictability stimuli in the LIFG in relation to standardized age of acquisition.

Table 1. *Mean (SD) for demographic variables, and language and standardized measures for the three groups of participants. Groups did not differ on any of the variables measured with the exception of age of second language acquisition and self-reported second language proficiency (p-values indicated in bold font represent significant group differences).*

	Simultaneous (n=10; 3 males; 8 L1 English)	Early (n=12; 1 male; 8 L1 English)	Late (n=8; 3 males; 7 L1 English)	*p* (one-way ANOVA)
Age	23.1 *(2.9)*	24.8 *(3.9)*	26.6 *(3.7)*	.13
Education	15.4 *(1.8)*	15.2 *(1.8)*	16 *(1.9)*	.64
AoA (years)	0 *(0)*	4.7 *(0.6)*	6.9 *(1.1)*	**<.01**
L1 letter fluency (letters F, A, and S)	39.8 *(14.2)*	41.2 *(9.1)*	41 *(15.4)*	.97
L2 letter fluency (letters P, F, and L)	34 *(9.1)*	27.1 *(10.1)*	26.9 *(10.9)*	.21
L1 category fluency (animals)	24.4 *(7.5)*	21.0 *(6.8)*	26 *(8.1)*	.31
L2 category fluency (fruits)	17.1 *(3.7)*	15.7 *(3.4)*	13.5 *(4.3)*	.15
Digit Span Forward[a]	7.5 *(1.2)*	7.1 *(1.2)*	7.1 *(1.2)*	.70
Digit Span Backward[b]	5.3 *(1.4)*	5.5 *(1.1)*	5.6 *(1.4)*	.86
Digit Span Sequencing[c]	6.3 *(1.2)*	6.5 *(1.4)*	6.5 *(1.4)*	.93
Letter Number Sequencing[d]	6.3 *(1.4)*	6.0 *(1.3)*	6.0 *(1.3)*	.85
Matrix Reasoning[e]	20.3 *(2.9)*	19.0 *(3.1)*	19.4 *(5.0)*	.70

AoA=age of second language acquisition; L1=native language; L2=Second language

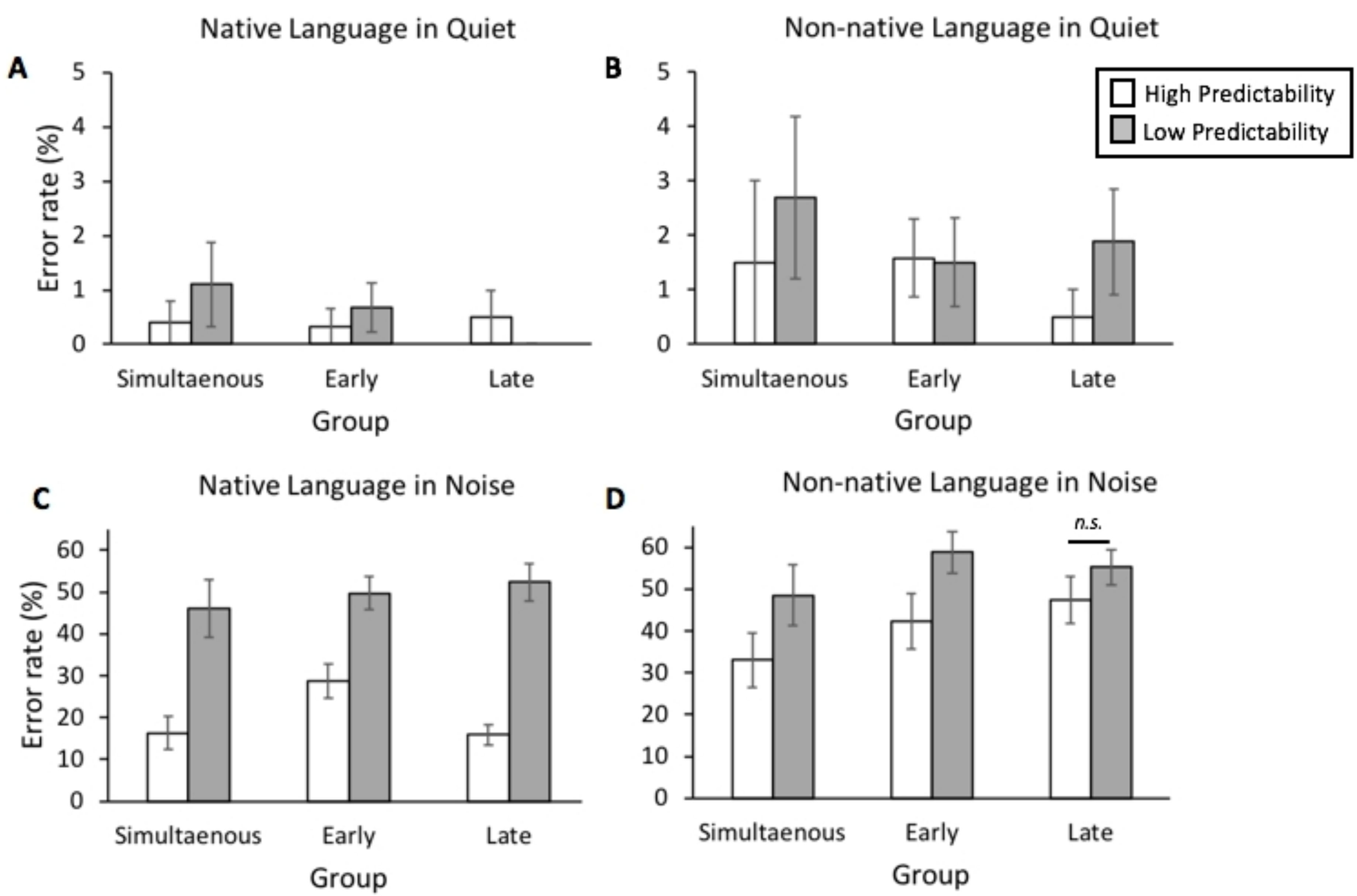
[a] standardized scores; maximum 9
[b] standardized scores; maximum 8
[c] standardized scores; maximum 9
[d] standardized scores; maximum 8
[e] maximum 26

Table 2. *Sample stimuli for the speech perception task.*

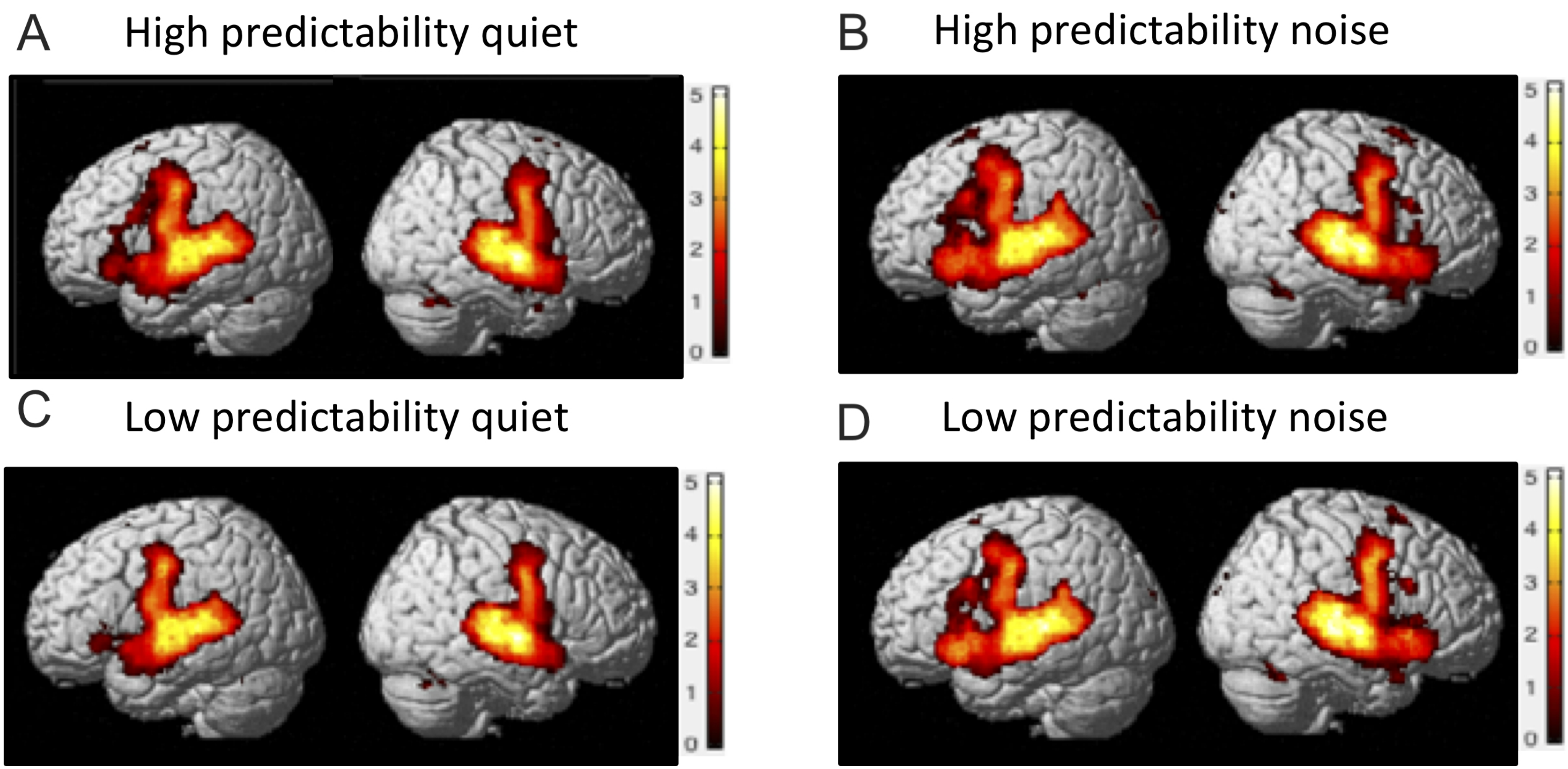
	High predictability	**Low Predictability**
English	Stir your coffee with a spoon.	Bob could have known about the spoon.
French	Marie portait ses cheveux en tresses.	Tu ne devrais pas parler des tresses.

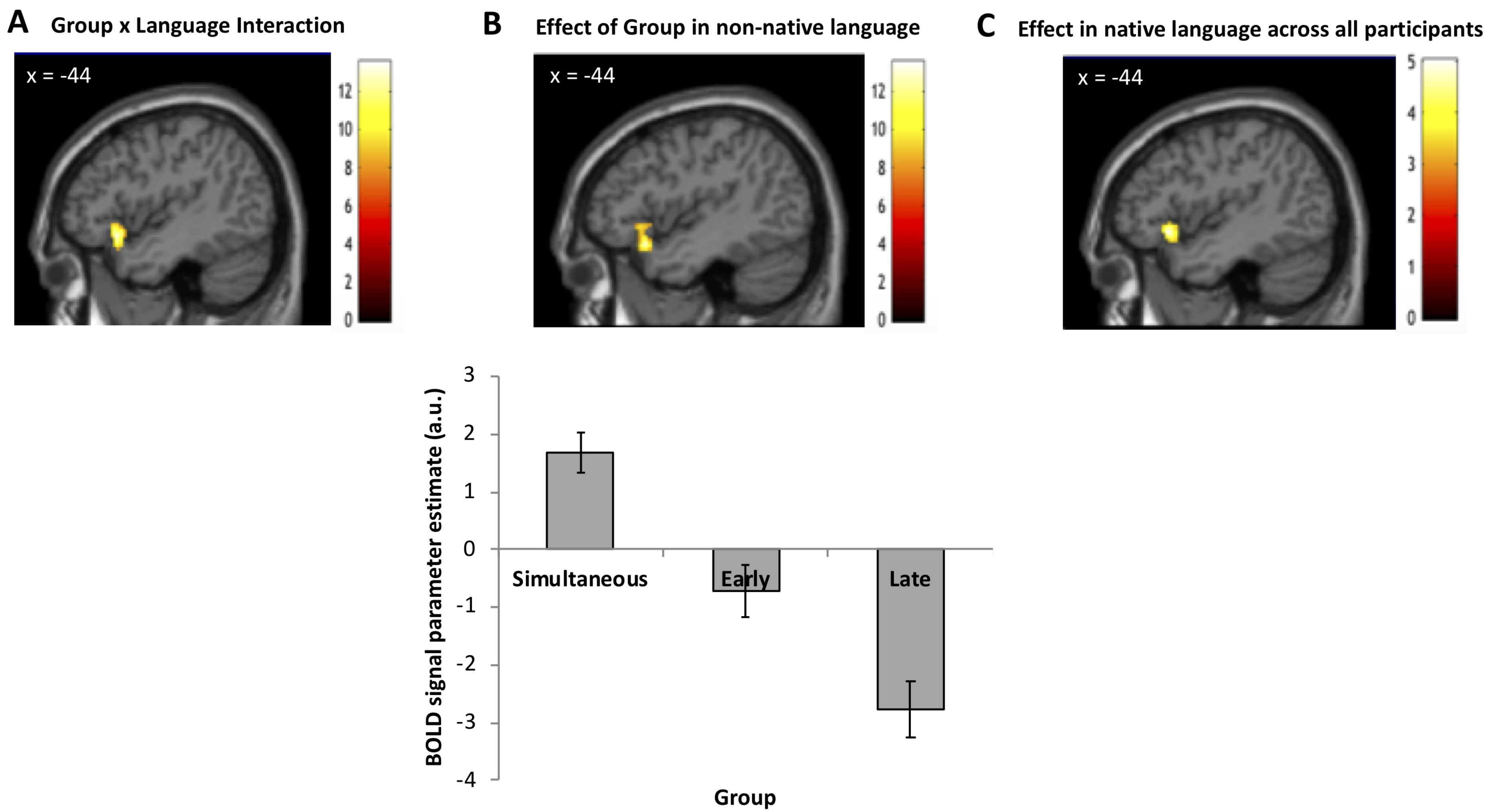
Table 3. *Results of the fMRI analyses. Peak voxels, cluster extents, and Z-scores are shown for each analysis. All reported results were obtained with an uncorrected p<.001and cluster extent threshold of 100 voxels, and survived cluster-level correction at p(FDR)<.05.*

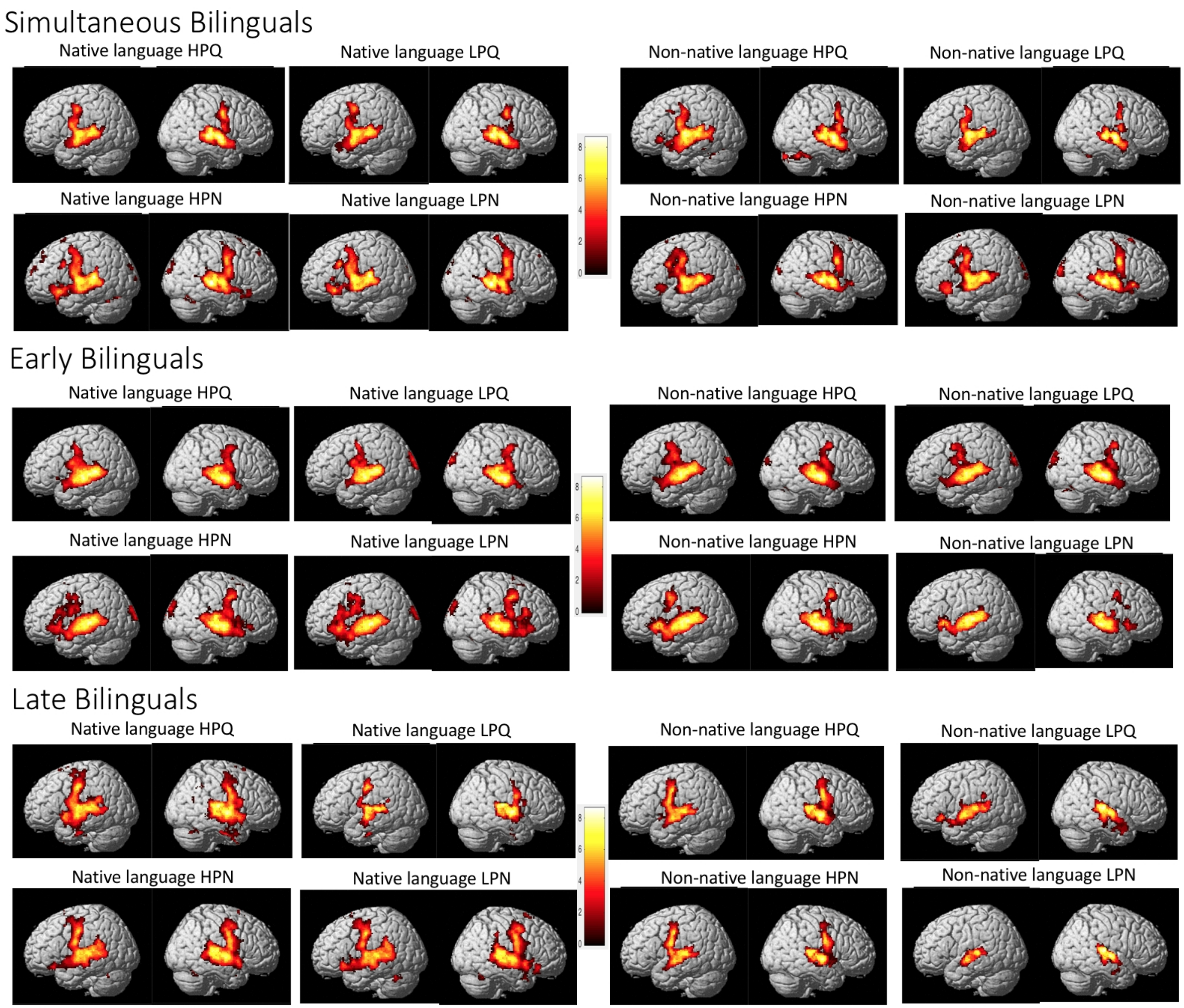
Brain region	Cluster extent (voxels)	MNI coordinates (x, y, z mm)	Z-score
LPN-HPN ***Language x Group interaction***			
Left inferior frontal gyrus	183	-46 20 -12	4.12
Effect of Group in non-native language			
Left inferior frontal gyrus	207	-42 22 -18	4.14
LPN-HPN in native language			
Left inferior frontal gyrus	130	-44 22 -8	4.54



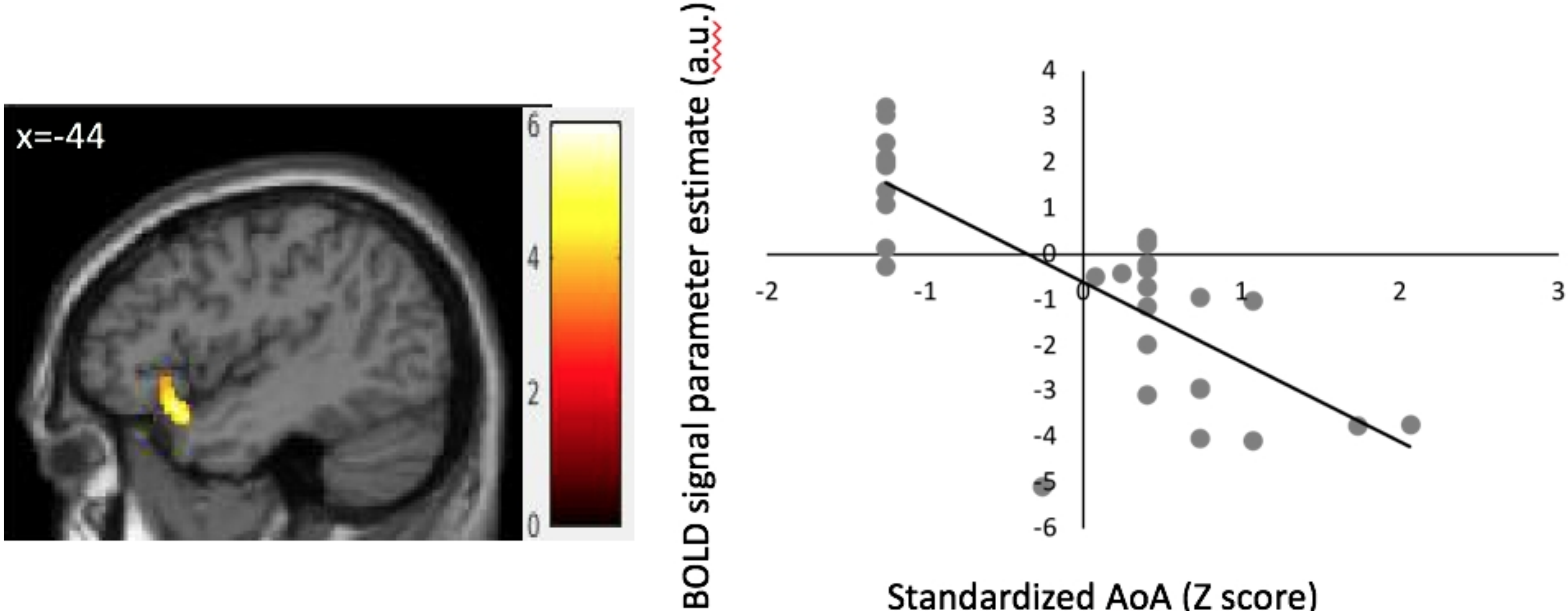
A
Native Language in Quiet
Error rate (%)
0
1
2
3
4
5
Simultaenous
Early
Late
Group
B
Non-native Language in Quiet
0
1
2
3
4
5
High Predictability
Low Predictability
Simultaenous
Early
Late
Group
C
Native Language in Noise
Error rate (%)
0
10
20
30
40
50
60
Simultaenous
Early
Late
Group
D
Non-native Language in Noise
0
10
20
30
40
50
60
n.s.
Simultaenous
Early
Late
Group


A
High predictability quiet
B
High predictability noise
C
Low predictability quiet
D
Low predictability noise
5
4
3
2
1
0


A Group x Language Interaction
B Effect of Group in non-native language
C Effect in native language across all participants
x = -44
x = -44
x = -44
0 2 4 6 8 10 12
0 2 4 6 8 10 12
0 1 2 3 4 5
BOLD signal parameter estimate (a.u.)
3 2 1 0 -1 -2 -3 -4
Simultaneous
Early
Late
Group



HPQ = high predictability in quiet; HPN = high predictability in noise; LPQ = low predictability in quiet; LPN = low predictability in noise


x=-44
6
4
2
0
BOLD signal parameter estimate (a.u.)
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-2
-1
0
1
2
3
Standardized AoA (Z score)

Declarations of interest

The authors declare no conflict of interest.