EJN EUROPEAN JOURNAL OF NEUROSCIENCE

FENS Federation of European Neuroscience Societies

European Journal of Neuroscience, Vol. 40, pp. 3387–3395, 2014 doi:10.1111/ejn.12673

COGNITIVE NEUROSCIENCE

When first language is not first: an functional magnetic resonance imaging investigation of the neural basis of diglossia in Arabic

Michael Nevat,[1,2] Asaid Khateb[1,2,3] and Anat Prior[2,3]

[1]The Unit for the Study of Arabic Language, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Faculty of Education, University of Haifa, Mount Carmel, Haifa 31905, Israel

[2]The Laboratory for the Study of Bilingualism, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Faculty of Education, University of Haifa, Haifa, Israel

[3]Department of Learning Disabilities, Faculty of Education, University of Haifa, Haifa, Israel

Keywords: Arabic, bilingualism, healthy subjects, Hebrew, semantic categorization

Abstract

In Arabic, the language used for everyday conversation ('spoken Arabic' – SA) differs markedly from literary Arabic (LA), which is used for written communication and formal functions. This fact raises questions regarding the cognitive status of the two varieties and their processing in the brain. Previous studies using auditory stimuli suggested that LA is processed by Arabic native speakers as a second language. The current study examined this issue in the visual modality. Functional magnetic resonance imaging (fMRI) responses were collected while Arabic–Hebrew bilinguals performed a semantic categorization task on visually presented words in LA, SA and Hebrew. Performance on LA was better than SA and Hebrew, which did not differ from each other. Activation in SA was stronger than in LA in left inferior frontal, precentral, parietal and occipito-temporal regions, and stronger than in Hebrew in left precentral and parietal regions. Activation in SA was also less lateralized than activation for LA and Hebrew, which did not differ from each other in terms of lateralization, though activation for Hebrew was more extensive in both hemispheres than activation for LA. Altogether, these results indicate an advantage for LA in the current study, presumably due to participants' proficiency in reading in this language. Stronger activation for SA appears to be due to the relative unfamiliarity of written word forms in SA, which could also explain differences in performance between the two languages. However, the stronger activation observed in the left parietal cortex may also reflect stronger associations among words in SA.

Introduction

The Arabic language is a typical example of 'diglossia' (Ferguson, 1959), which is a socio-linguistic situation in which the language used for everyday conversation (i.e. spoken Arabic – SA), differs markedly from the written language (i.e. literary Arabic – LA, also called modern standard Arabic). SA and LA varieties differ in terms of the age and manner of acquisition as well as their use. SA, the spoken local dialect, is the first language (L1) acquired by native speakers of Arabic, serves strictly for oral communication and does not typically exist in written form. LA, a highly codified form, is acquired later in childhood primarily through formal education, though children are differentially exposed to it aurally through audio-visual media (Abu-Rabia, 2000; Saiegh-Haddad, 2003; Boudelaa & Marslen-Wilson, 2013). LA is used for reading and writing and formal speech functions (religious sermons, official speeches, news broadcasts and teaching; Ibrahim & Aharon-Peretz, 2005; Levin *et al.*, 2008; Saiegh-Haddad *et al.*, 2011). LA differs from SA in the phonological, morpho-syntactic and lexical-semantic domains (Saiegh-Haddad *et al.*, 2011). Also, whereas LA is homogeneous across the Arabic-speaking world, considerable differences exist between dialects of SA used in different regions (Ayari, 1996; Saiegh-Haddad, 2003, 2005; Saiegh-Haddad *et al.*, 2011; Boudelaa & Marslen-Wilson, 2013).

Implications of diglossia for the acquisition of basic reading processes by Arabic speakers have repeatedly been discussed (Saiegh-Haddad *et al.*, 2011), and the extent to which it resembles bilingualism has been examined only behaviorally. For instance, Eviatar & Ibrahim (2000) reported that the performance of Arabic-speaking children on tests assessing meta-linguistic abilities resembled that of bilinguals. Also, Ibrahim & Aharon-Peretz (2005) investigated the cognitive status of SA and LA (and Hebrew as a second language, L2) using semantic priming effects during auditory lexical decision. They reported that priming effects were larger when primes were in SA and target words were either in LA or in Hebrew than the reverse (LA or Hebrew primes and SA targets). Furthermore, the priming effects for LA and Hebrew were identical (Ibrahim, 2009). These effects resemble findings in bilinguals (Keatley *et al.*, 1994;

Correspondence: Prof A. Khateb, [1]The Unit for the Study of Arabic Language, as above.
E-mail: akhateb@edu.haifa.ac.il

Received 7 April 2014, revised 9 June 2014, accepted 17 June 2014

Gollan *et al.*, 1997), which show larger forward priming (from L1 to L2) than backward priming (from L2 to L1), and supported the view that SA is cognitively represented as L1 and LA as L2 (Ibrahim & Aharon-Peretz, 2005).

The current study assessed the neural basis of diglossia by analysing the processing of visually presented LA and SA words in adult Arabic speakers, and comparing both to the participants' formal L2 (Hebrew). Participants performed a semantic categorization task, previously shown to reliably activate left hemisphere language areas (Seghier *et al.*, 2004, 2008). Functional magnetic resonance imaging (fMRI) and behavioral measures were analysed to investigate whether separate neural processes might support visual word processing in the languages examined.

Because the diglossic situation is unique in that it distinguishes between an L1 for oral communication and a different L1 for literacy, stronger activation could be expected in areas differentiating the weaker from the dominant language (see examples in Wartenburger *et al.*, 2003; Meschyan & Hernandez, 2006; Rüschemeyer *et al.*, 2006) when comparing visual word processing in SA or Hebrew to LA, with a performance advantage for LA.

Materials and methods

Participants

Twenty-six healthy women participated in the study. Due to a technical malfunction, the scanning of one participant could not be completed and she was therefore excluded from the analyses. The remaining participants' age was between 18 years and 4 months and 23 years and 10 months (mean = 20.3 years, SD = 1.4 years). All participants were Arabic–Hebrew bilinguals, in their first year of undergraduate studies in the Faculty of Education at the University of Haifa. Participation in the experiment counted as partial fulfillment of course credit. All participants were right-handed except for one [handedness laterality index (LI) mean = 58.6, SD = 18.7], had normal or corrected-to-normal vision, had no history of neurological or psychiatric disorders, and reported no use of any psychoactive medication at the time of the experiment. Participants were informed about the purposes of the study and gave written informed consent before participating. The study protocol was approved by the ethics committee of the University of Haifa and conforms with the World Medical Association Declaration of Helsinki.

Concerning the participants' language background, they had all been exposed to formal instruction of LA since the first grade, and of Hebrew as an L2 since the third grade (at the age of 8–9 years). They had all completed their secondary school studies and successfully passed the high-school matriculation exams with Hebrew as L2. In addition, all students were required to pass the Hebrew Proficiency Test ('YAEL' – National Institute for Testing & Evaluation) with a score of at least 115 (out of 150 possible points) as a prerequisite for admission to the University, where they followed their courses in Hebrew. Therefore, at the time of the experiment, all could be considered as proficient bilinguals.

Stimuli and procedure

Participants performed a semantic categorization task based on Seghier *et al.* (2004), in which they were asked to judge whether pairs of words presented on the screen were semantically related (SR; i.e. belonged to the same semantic category) or not. The stimuli were high-frequency, concrete imageable nouns, in Arabic and Hebrew. For the selection of the stimuli, a questionnaire was first presented to a group of 30 native Arabic-speaking participants, who were asked to rate the frequency of 220 SA words and 220 LA words using a scale from 0 to 6 (0 unknown or least frequent, 6 most frequent). All stimuli were selected so as to minimize phonological overlap between SA and LA words (but also with Hebrew translation equivalent). The most frequent 200 words in each language variety (LA – mean = 4.34, SD = 0.95; SA – mean = 4.75, SD = 1.17) were then retained for this study.

From the selected words (length three-six letters), 50 SR word pairs and 50 semantically unrelated (SU) word pairs were formed in each language (SA and LA). These pairs were then presented in a questionnaire to another group of Arabic-speaking participants, who rated semantic relatedness of the words in each pair using a scale from 0 to 5 (0 for least related and 5 for most related) to ensure their suitability for the different conditions. The average relatedness in each list was above 3 for related pairs (LA = 3.87 $\pm$ 0.38; SA = 3.60 $\pm$ 0.18), and below 0.5 for unrelated pairs (LA = 0.28 $\pm$ 0.18; SA = 0.28 $\pm$ 0.17). Finally, the readability of these pairs was assessed in a pilot study conducted with 10 adult participants using a computerized speeded semantic judgment task (as in Khateb *et al.*, 2003). An item-by-item analysis performed on the results of this study allowed selecting only pairs that yielded at least 8/10 correct responses. For the Hebrew words, the stimuli consisted of translation equivalents of words from SA and LA.

The words in each pair were simultaneously displayed, one word beneath the other, at the center of a computer screen located outside the scanner. Mirrors fastened to a head coil reflected the stimuli, so that they could be viewed by the participants. We used a block paradigm that alternated between the semantic categorization and the control condition. In the control condition, pairs of Greek character strings were simultaneously presented, and participants judged whether the strings were physically identical or not. In this condition, which was used for all language blocs and mainly involves visual processing, the stimuli were designed so as to resemble those in the semantic categorization blocks in terms of the number of characters and the spatial extent.

The same semantic categorization task was used for the three languages: one using stimuli in LA; one using stimuli in SA written in the Arabic orthography; and one using stimuli in Hebrew. Because the participants were all skilled readers in Arabic and Hebrew, the words in all language blocs were presented without diacritics (i.e. short vowels) as is customary for adult readers (Abu-Rabia, 2001). The order of presentation of the language runs was counterbalanced across participants. In each of the runs the activation condition consisted of 72 word pairs (48 SR pairs and 24 SU pairs) divided into six blocks. The subjects performed a yes/no task, and gave a response (using their left thumb) to indicate whether the two words in each pair were related or unrelated. The control condition also consisted of 48 pairs of identical Greek letter-strings and 24 pairs of visually different strings. Here, participants indicated whether the two strings in each pair were visually identical, or not. In all conditions (both activation and control) and language runs, stimulus pairs were presented every 2 s on the screen for 600 ms and in blocks of 24 s, repeated six times per condition. Hence, alternating blocks of activation–control conditions yielded a total duration of 4.8 min per language run. Responses were given using an MR-compatible response box that allowed registering the performance of the subject and the reaction times (RTs). In order to ensure full comprehension of the task demands, participants were provided with instructions before entering the scanner and underwent a training session of a few trials.

fMRI acquisition

The experiments were conducted using a 3T MRI scanner (GE Discovery MR750) at the Rambam medical center in Haifa. A high-resolution T1-weighted anatomical scan was recorded for each participant [voxel size – 1 × 1 × 1 mm; number of slices – 148; repetition time (TR) = 12.73 ms; echo time (TE) = 5.42 ms]. For each experimental run (task), 145 dynamic volumes with axial contiguous ascending acquisitions were recorded (voxel size – 3.44 × 3.44 × 3.4; matrix size – 64 × 64; number of slices – 43; interslice gap – 0%; TR = 2000 ms; TE = 30 ms; field of view – 220; flip angle = 60°). For each run, the functional scanning was always preceded by 10 s of dummy scans to insure tissue steady-state magnetization.

Whole-brain analysis

MRI data were analysed with the Statistical Parametric Mapping SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/). All functional volumes were subjected to standard preprocessing procedures (Friston *et al.*, 2007), including: spatial realignment; normalization [to Montreal Neurological Institute (MNI) space with 2 × 2 × 2 mm^3 voxel size]; and smoothing with an isotropic 5-mm full-width at half-maximum Gaussian kernel. Time-series from each voxel were high-pass filtered (1/128 Hz cutoff). After preprocessing, analyses at the level of the individual participant (first-level analysis) were performed using the general linear model applied to each voxel (Friston *et al.*, 1995; Worsley & Friston, 1995) and an auto-regressive [AR (1)] function to account for temporal correlations between them across the whole brain. Each run was modeled as a distinct session, and each condition within a run (semantic categorization or control) was separately modeled. At the group level, a two-way (language × condition) 'flexible factorial' model was then specified, resulting in a total of six regressors (three languages × two conditions).

Activation in all semantic categorization blocks was compared with activation in all control blocks in order to identify all regions that were active during processing of stimuli in at least one of the languages. Afterwards, activation during semantic categorization in each of the languages was compared with activation during control categorization in the same run. The resulting differences were then entered into a conjunction analysis in order to identify regions active in all languages, and were then compared with each other in order to examine differences in activation between languages. In these latter comparisons, the results were masked with the activation map obtained for each language minus its baseline, to ensure that the resulting differences were due to activation rather than deactivation. All results reported here were obtained by specifying a threshold of $P_{FWE} < 0.05$, and minimal cluster extent of 10 voxels. Data inspection before the second-level analysis revealed that one of the participants exhibited right hemisphere dominance, and she was therefore excluded from further analyses.

Lateralization of brain activation

The semantic categorization task presented to participants in the current study has been found to yield highly lateralized activation when the task was presented in participants' L1 (Seghier *et al.*, 2004). Performing semantic judgment tasks in L2 has been found to recruit regions in the right lateral frontal cortex, such as the precentral gyrus (Rüschemeyer *et al.*, 2006) and middle frontal gyrus (MFG; Wartenburger *et al.*, 2003), when compared with L1. This could be expected to result in reduced lateralization for L2. In order to determine whether lateralization of activation varied between languages in the current study, LIs were calculated for each participant in each language, as described by Seghier *et al.* (2004). The following formula was used to calculate LIs:

$$\mathrm{LI} = (\mathrm{voxels_{left}} - \mathrm{voxels_{right}})/(\mathrm{voxels_{left}} + \mathrm{voxels_{right}})$$

where $\mathrm{voxels_{left}}$ and $\mathrm{voxels_{right}}$ are the number of voxels in a given map exceeding a selected significance threshold, residing in the left and right hemispheres, respectively. The threshold used was $P_{unc.} < 0.005$, as in Seghier *et al.* (2004). LIs thus calculated were entered into a repeated-measures ANOVA to examine effects of language.

Regions of interest (ROIs) analyses

ROIs were defined in order to further examine effects of language. For this purpose, areas active during semantic categorization were identified based on the comparison between semantic categorization and control across languages. Such a comparison yields regions active during categorization in any (or all) of the languages, and therefore should not be expected to bias (as per the concerns raised by Kriegeskorte *et al.*, 2009) comparisons among languages, which were the objective of these analyses. The volume of the ROIs was thresholded at 500 mm^3. Due to extensive activation in the left frontal cortex, the activated voxels in this region were divided into six regions, based on the automated anatomical labeling atlas (AAL; Tzourio-Mazoyer *et al.*, 2002). In total, 11 ROIs were examined: (i) left triangularis; (ii) left opercularis; (iii) left MFG; (iv) left insula; (v) left precentral gyrus; (vi) left postcentral gyrus. Additional ROIs were defined in the: (vii) left middle temporal gyrus (MTG); (viii) left parietal (on the border between the parietal and occipital lobes); (ix) supplementary motor area (SMA); (x) right insula; and (xi) left fusiform gyrus.

Regarding the left fusiform ROI, this was defined based on examination of images of the differences between categorization and control blocks generated for each of the participants separately. A threshold of $P_{unc.} < 0.001$ was applied, with a cluster extent of 75 voxels. This was done in order to reduce the probability that voxels that were strongly activated only in a minority of participants would be included in the ROI, thereby ensuring the correct identification of region(s) selectively involved in the processing of written words among a majority of participants. The resulting images were then masked so as to include only active voxels residing either in the left inferior temporal gyrus, the left fusiform gyrus or the left inferior occipital gyrus according to the AAL. The average of the masked and thresholded images was calculated, and a threshold value of 4.87 (corresponding to $P_{FWE} \approx 0.05$) was applied to this image.

Peri-stimulus time histograms were obtained for each participant, in each region, in each language, using the MarsBaR toolbox for SPM (v0.43; Brett *et al.*, 2002). The measure of brain activity selected was the mean signal change over the interval between 6 and 30 s from the onset of blocks. These measures were used to examine effects of language in the ROIs, as well as correlations between activations in regions exhibiting effects of language.

Behavioral analysis

The individual mean RTs were computed from trials in which correct responses were recorded, and were analysed in a two-way repeated-measures ANOVA with language and condition (semantic categorization vs. control) as within-subject factors. A similar analysis was performed on the individual accuracy rates computed separately for each condition and language. All analyses were performed using IBM SPSS Statistics software (v.19.0). In all ANOVAS, the Greenhouse–Geisser correction was applied for sphericity values lower than 0.75, and the Huynh–Feldt correction was applied for sphericity values greater than 0.75 (Field, 2005).

Results

Behavioral measures

Accuracy

The analysis of accuracy was computed as the percentage of correct responses relative to the number of trials where a response was obtained (Table 1). The two-way repeated-measures ANOVA performed on the individual values of accuracy showed that performance was higher in the control than in the experimental (semantic categorization) condition ($F_{1,24} = 59.3$, $P < 0.00001$). There was also a significant main effect of language ($F_{2,48} = 20.1$, $P < 0.00001$) due to the fact that accuracy was higher in LA than in SA (Fisher's LSD *post hoc* tests, $P < 0.00001$) and in Hebrew ($P < 0.00001$). The two-way interaction between language and task was also significant ($F_{2,48} = 24.7$, $P < 0.00001$). This interaction was due to the fact that the language effect was significant only in the activation ($F_{2,48} = 26.3$, $P < 0.00001$), but not in control blocks ($P = 0.83$; Table 1).

TABLE 1. Accuracy and RTs on semantic categorization and control tasks, by language

	Mean accuracy in % (±SD)	Mean RT in ms (±SD)
LA activation	90.5 (7.8)	954 (131)
SA activation	79.6 (10.8)	1074 (140)
Hebrew activation	79.3 (7.9)	1068 (219)
LA control	93.5 (4.1)	667 (116)
SA control	93.6 (5.0)	656 (156)
Hebrew control	93.9 (4.5)	666 (145)

LA, literary Arabic; RT, reaction time; SA, spoken Arabic.

RTs

Mean RTs (±SD), based on trials with correct responses in each condition and language are presented in Table 1 (right column). The two-way repeated-measures ANOVA performed on the individual RTs using condition and language as within-subject factors showed a highly significant main effect of condition due to shorter responses in control (mean = 663 ms) than in activation blocks (mean = 1032 ms, $F_{1,24} = 266.9$, $P < 0.00001$), and a main effect of language ($F_{2,48} = 6.1$, $P < 0.005$) due to faster responses in LA than in SA ($P < 0.005$) and in Hebrew ($P < 0.004$), which did not differ from each other. The interaction between the two factors was also significant ($F_{2,48} = 9.7$, $P < 0.0003$), due to the fact that the condition effect (i.e. difference between activation and control) was slightly smaller in LA than in the two other languages (Table 1).

fMRI analysis

Whole-brain analysis

The activation maps for the comparison between activation and control conditions, analysed in all languages together, are displayed in Fig. 1, which shows a dominant left hemisphere activation pattern.

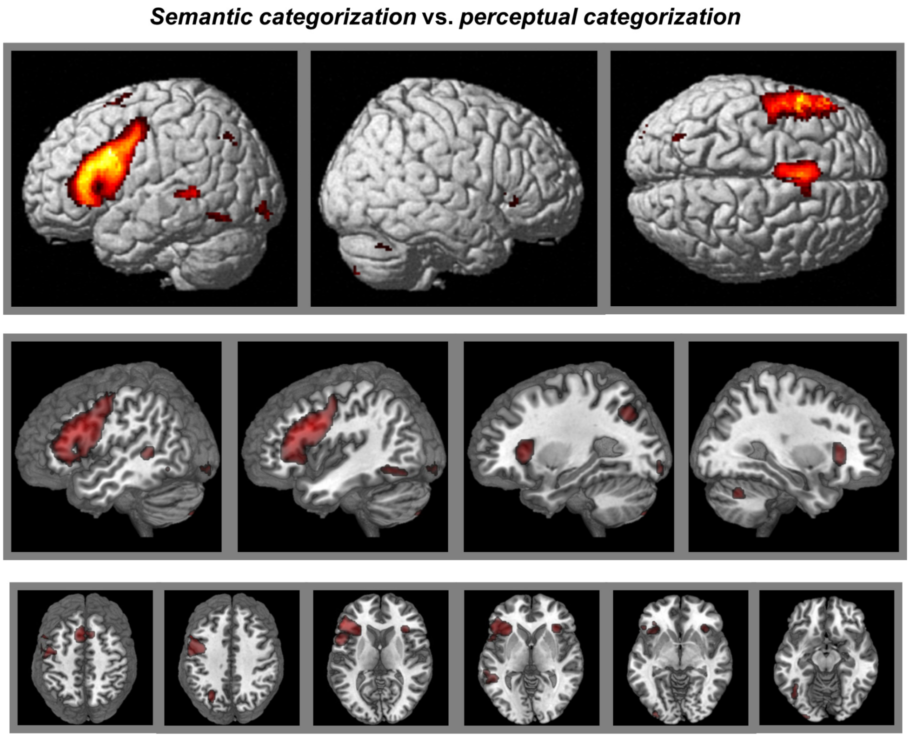

FIG. 1. Regions more activated during semantic categorization compared with control, across languages.

TABLE 2. Regions more active during semantic categorization compared with control, across languages

Semantic categorization vs. control				
Anatomical location (AAL)	BA	Coordinates X; Y; Z	Z-value	Cluster size
Left IFG pars triangularis	46	−44; 26; 22	12.22	3689
	46	−52; 32; 12	12.10	
	46	−44; 28; 14	11.50	
	9	−42; 14; 28	15.35	
Left IFG pars opercularis	44	−52; 14;6	7.65	
Left precentral gyrus	6	−44; 4; 30	12.61	
	4	−52; −4; 48	8.27	
Left insula	13	−32; 24; 4	8.80	
SMA	6	−4; 16; 56	10.14	776
	6	10; 18; 50	5.45	
Left inferior temporal & occipital gyri/fusiform gyrus	37	−44; −56; −12	6.68	55
		−44; −66; −14	5.45	
Left MTG	22	−52; −42; 4	6.08	151
		−60; −36; 4	5.66	
Left inferior parietal/middle occipital gyrus	7	−28; −66; 40	7.08	119
Left inferior & middle occipital gyri	18	−28; −96; −12	5.99	47
		−34; −92; −4	5.78	
Right insula/IFG (pars triangularis)		32; 26; 4	7.79	187
Right cerebellum		10; −78; −38	8.31	196

AAL, automated anatomical labeling; BA, Brodmann area; IFG, inferior frontal gyrus; MTG, middle temporal gyrus; SMA, supplementary motor area.

TABLE 3. Conjunction of regions activated in all languages, when comparing semantic categorization and control

Conjunction of categorization: LA, SA and Hebrew				
Anatomical location (AAL)	BA	X; Y; Z	Z	K
Left IFG/pars triangularis	9	−42; 14; 26	Inf	1693
	46	−52; 34; 12	Inf	
	46	−42; 26; 22	Inf	
	45	−50; 28; 16	Inf	
	13	−42; 30; 6	5.66	
Left insula	13	−34; 26; 4	5.38	
Left SMA	6	−4; 16; 58	6.18	169
Right cerebellum		12; −78; −38	5.29	17

AAL, automated anatomical labeling; BA, Brodmann area; IFG, inferior frontal gyrus; LA, literary Arabic; SA, spoken Arabic; SMA, supplementary motor area.

TABLE 4. Regions more active in SA compared with LA during semantic categorization

Semantic categorization: SA vs. LA				
Anatomical location (AAL)	BA	X; Y; Z	Z	K
IFG (pars opercularis)	9	−54; 8; 22	5.92	370
Precentral gyrus/postcentral gyrus	6	−50; 0; 38	6.19	
Inferior temporal gyrus	37	−50; −64; −8	5.21	15

AAL, automated anatomical labeling; BA, Brodmann area; IFG, inferior frontal gyrus; LA, literary Arabic; SA, spoken Arabic.

As detailed in Table 2, these activations included, antero-posteriorly in the left hemisphere, the inferior and middle frontal gyri (IFG and MFG, respectively), insula, precentral and postcentral gyri, the middle and inferior temporal gyrus, and the fusiform and inferior occipital gyri. In the right hemisphere, the activation included principally parts of the right SMA, superior frontal gyrus, insula, IFG and cerebellum. Figure 2A presents, based on activation vs. control in each language, a conjunction that shows the regions that were commonly activated in all languages. As displayed here and detailed in Table 3, the common activation was found almost exclusively in left frontal areas, including IFG, insula, precentral gyrus and MFG.

Six possible pairwise comparisons between languages were conducted [i.e. (i) LA vs. SA; (ii) LA vs. Hebrew; (iii) SA vs. LA; (iv) SA vs. Hebrew; (v) Hebrew vs. LA; (vi) Hebrew vs. SA]. Of these, only the comparison SA vs. LA yielded significant differences (at $P_{\mathrm{FWE}} < 0.05$) with stronger activation for SA than for LA (Fig. 2B; Table 4), though a threshold of $P_{\mathrm{unc.}} < 0.001$ did show differences between Hebrew and LA in the left precentral gyrus and medial frontal and left occipital cortices, and between SA and Hebrew in the left MTG and superior parietal lobule, and in the right inferior frontal cortex (precentral gyrus and IFG pars opercularis). Figure 2C presents a superposition of the commonly activated voxels in the conjunction map (red areas, as in A), and those exhibiting significant differences between SA and LA (superposed in blue, as in B). The regions that were more highly activated in SA relative to LA included the left IFG, precentral and postcentral gyri, and the left inferior temporal gyrus (Table 4).

LIs

The effect of language on LIs was significant ($F_{2,46} = 3.36$, $P < 0.05$). LI for LA was greater than for SA. LIs for Hebrew were also greater than for SA, though the difference was marginally significant. Examination of the number of supra-threshold voxels in each hemisphere yielded significant effects of language in both hemispheres ($F_{2,46} = 12.70$, $P < 0.001$ and $F_{2,46} = 10.39$, $P < 0.01$ for left and right hemispheres, respectively). In both hemispheres activation was most extensive in SA, followed by Hebrew and finally LA. However, in the left hemisphere differences between SA and LA were only marginally significant. Thus, it appears that whereas activation for Hebrew was proportionately more extensive in both hemispheres compared with LA, yielding similar values of LI, activation in SA was particularly extensive in the right hemisphere, resulting in lower values.

It is worthy of noting that LIs were generally lower than those found by Seghier *et al.* (2004, 2008), with values (mean ± SD) of 0.59 ± 0.22 for LA, 0.52 ± 0.21 for SA and 0.59 ± 0.2 for Hebrew. However, setting the threshold at $P_{\mathrm{FWE}} < 0.05$ resulted in LIs similar to those reported by Seghier *et al.* (2004, 2008), with 0.79 ± 0.21 for LA, 0.74 ± 0.2 for SA and 0.81 ± 0.19 for Hebrew. This may indicate that differences between LIs obtained in the two studies may be due to the fact that in the current study participants were scanned using a 3T scanner, whereas Seghier *et al.* (2004) used a 1.5T scanner, which may have resulted in more voxels, particularly in the right hemisphere, exceeding the threshold in the current study. On the other hand, lower LIs may have to do with the languages being examined; Al-Hamouri *et al.* (2005) report stronger activation in the right hemisphere for Arabic compared with Spanish, and attribute this difference to ambiguities in decoding written words in Arabic presented without vowel diacritics. This interpretation can be extended to unvowelled words in Hebrew, and may be a contributing factor to the lower LIs found across languages in the current study.

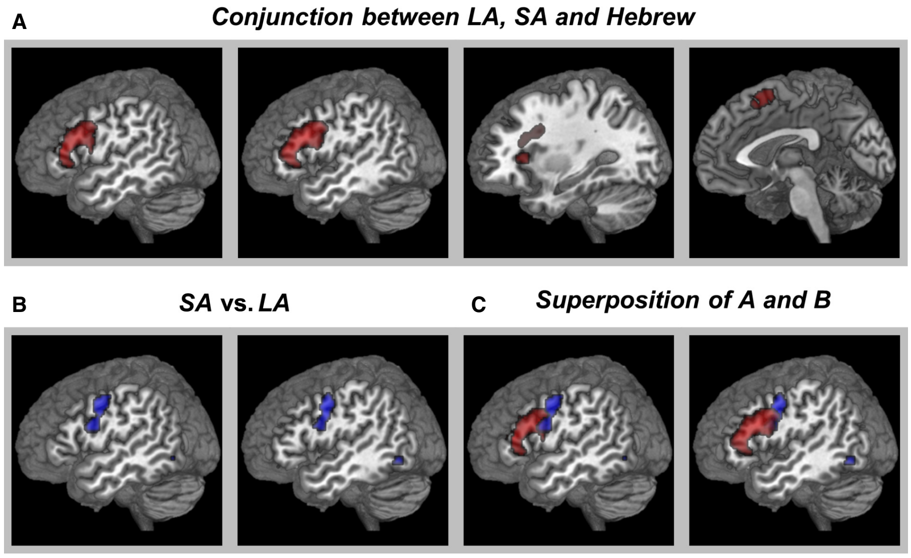

FIG. 2. (A) Regions activated in all three languages (red). (B) Regions activated more strongly in spoken Arabic (SA) compared with literary Arabic (LA; blue). (C) Superposition of (A) and (B).

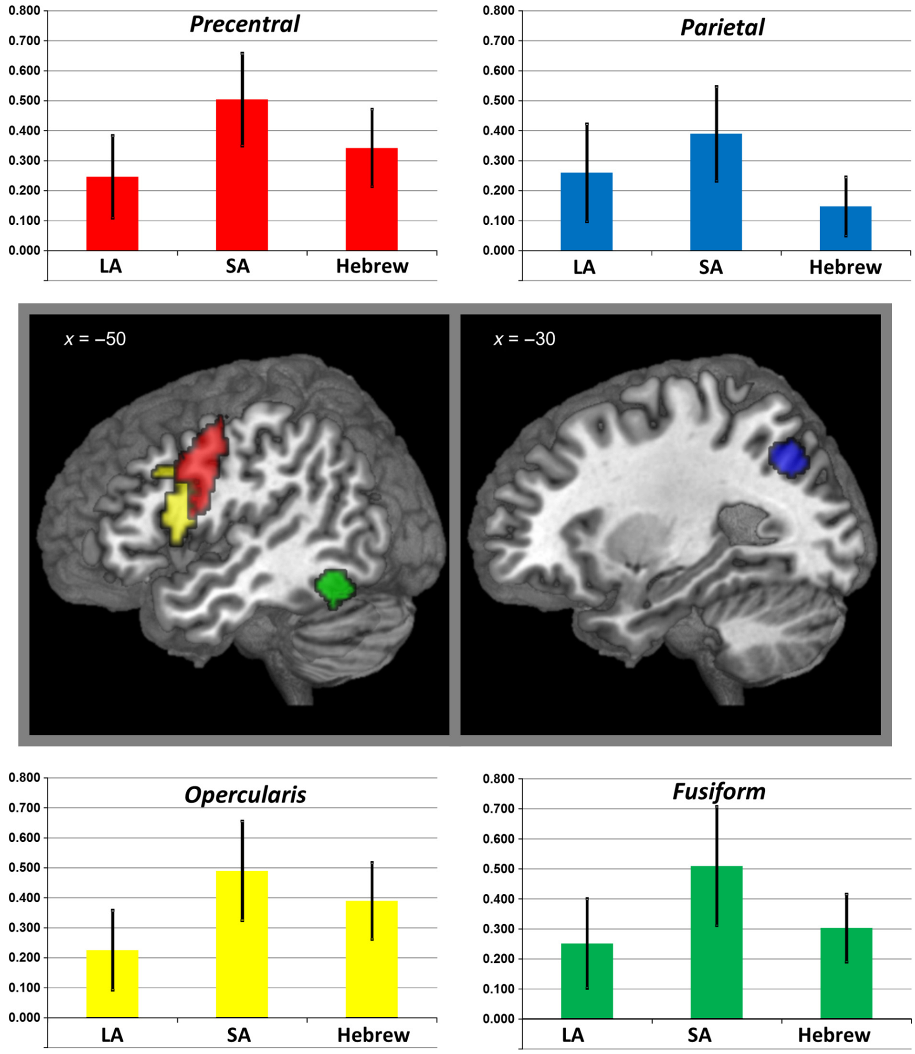

FIG. 3. Regions of interest (ROIs) exhibiting significant effects of language. Error bars indicate limits of 95% confidence intervals about the mean.

ROI analyses

Significant effects of language were found in four regions: left opercularis ($F_{2,46}$ = 4.7, $P < 0.05$); left precentral ($F_{2,46}$ = 4.16, $P < 0.05$); left parietal ($F_{2,46}$ = 4.46, $P < 0.05$); and left fusiform ($F_{2,46}$ = 3.51, $P < 0.05$). Results of pairwise comparisons indicated that activation for SA was stronger than for LA in all four regions, whereas activation for Hebrew did not differ significantly from activation for LA. Furthermore, in two of the four regions (left precentral and left parietal), activation for SA was also stronger than activation for Hebrew, and the difference between SA and Hebrew was marginally significant in the left fusiform gyrus (Fig. 3). Marginally significant effects of language were also found in the left triangularis ($F_{2,46}$ = 2.67, P = 0.081, SA > LA) and postcentral gyrus ($F_{2,46}$ = 3.02, P = 0.061, SA > LA). Additional analyses of the regions, which best differentiated SA from LA and Hebrew, showed that the activity during processing of SA yielded significant correlations between the precentral gyrus and the left parietal, and between the precentral and the fusiform gyrus (r = 0.68, $P < 0.01$; r = 0.63, $P < 0.01$, respectively, after correction for multiple comparisons). It should be noted that the differences found here between SA and LA in the parietal lobe had not been evident in the whole-brain analysis. However, when in this analysis the threshold was lowered to $P_{\text{unc.}} < 0.001$ (instead of $P_{\text{FWE}} < 0.05$), a significant cluster did indeed emerge.

Discussion

In the current study, our objective was to compare the processing of visually presented SA and LA words in Arabic–Hebrew bilinguals. For this purpose, participants performed a semantic categorization task on written words in LA, SA and Hebrew. We predicted that the processing of SA and LA words would reflect their history of acquisition and patterns of use. Thus, we expected that SA, while being the first language acquired by Arabic speakers but usually not encountered in the written form, would show response and activation patterns mimicking either L2 words or low-frequency/unfamiliar words. In contrast, because LA is the first written form acquired and more frequently used in writing, we predicted superior behavioral performance for LA words relying on the classical left language network.

Our results showed that accuracy was high for all languages, confirming the ability of the participants to correctly identify the stimuli. As predicted, categorization of LA words was faster and more accurate than categorization of words in either SA or Hebrew, which did not differ from each other.

In terms of brain activity, whole-brain analysis first showed that categorization vs. control conditions across all three language varieties revealed a classical pattern of activation consisting mainly of left hemisphere areas (Seghier *et al.*, 2004, 2008). Conjunction analysis between the three language varieties showed a strong involvement in the left hemisphere of the IFG, MFG, insula and SMA. When contrasting the three varieties, differences were observed only between SA vs. LA, with stronger activation for SA in left frontal and temporal areas.

We followed this whole-brain analysis with an ROI analysis, in which ROIs were selected based on areas most strongly distinguishing between categorization and control across languages. In this analysis, a more complex pattern emerged. SA generated stronger activation than LA in four of the regions examined in the left hemisphere, namely opercularis, precentral, parietal and fusiform. In addition, SA generated stronger activation than Hebrew in left precentral and left parietal areas. In all of these regions, activation for LA did not differ from activation for Hebrew.

Lateralization of activation was examined using LIs. In these analyses, lateralization for LA was stronger than for SA, and lateralization for Hebrew was similar to that found for LA, though only marginally stronger than for SA. This pattern was due to differences in the extent of activated voxels in each hemisphere in each of the languages. Activation in both hemispheres, but particularly in the right one, was most extensive for SA. Activation for Hebrew was less extensive than for SA, but more extensive than for LA. These results were supported by the fact that when a more lenient threshold was applied to whole-brain analyses, differences between Hebrew and LA, and between SA and Hebrew emerged, including differences in the right inferior frontal cortex.

Results of the current study reflect participants' higher proficiency in LA compared with Hebrew. This is particularly evident in participants' performance but, also, albeit more subtly, in measures of brain activation. Thus, comparisons between Hebrew and LA using a threshold of $P_{\text{unc.}} < 0.001$ revealed differences in the left precentral gyrus and in the SMA, while counting active voxels in each hemisphere reveals more extensive activation bilaterally for Hebrew.

Stronger activation in the bilateral precentral gyri and SMA has been reported during reading of sentences in L2 compared with L1, in the context of semantic and grammatical acceptability judgment tasks (Rüschemeyer *et al.*, 2006). Stronger activation in the SMA has also been found during reading of words in a less proficient L1 compared with a more proficient L2 (Meschyan & Hernandez, 2006). Finally, Wartenburger *et al.* (2003) report proficiency-related differences in activation in the left IFG during a semantic judgment task, at a location adjacent to the foci of differences in the precentral gyrus in the comparison between Hebrew and LA.

Regarding the observed differences between SA and LA in terms of performance and brain activity, these may be related to the relative unfamiliarity of written words in SA. Given the fact that words in SA are less often encountered in the written form than words in LA, effects of familiarity for SA word forms are to be expected in visual presentation. Such differences would most likely emerge in the left fusiform gyrus, where word frequency has been found to affect activation during reading (Joubert *et al.*, 2004; Kronbichler *et al.*, 2004). In this context, differences in lateralization of activation between SA and LA, which are related to more extensive activation in the right inferior frontal cortex for SA, may indicate that participants had to effortfully avoid interpreting words in SA as words in LA. Previous studies reported in the involvement of the right IFG during increased processing demands and inhibition during 'go/no-go' tasks (Chikazoe *et al.*, 2007; Lenartowicz *et al.*, 2011). Additionally, presentation of written words in SA may have resulted in enhanced sub-lexical phonological processing. Evidence of such processes during reading of written words in SA has been previously presented by Bentin & Ibrahim (1996).

Results concerning pseudoword reading are of relevance to the investigation of sub-lexical phonological processing, as reading of pseudowords is thought to require reliance on such processing. Greater activation for pseudowords compared with words has been reported in the left IFG during overt (Carreiras *et al.*, 2007; Heim *et al.*, 2013) and silent (Joubert *et al.*, 2004) reading. Activation in the left parietal cortex [intra-parietal sulcus (IPS)] has been reported during both silent and overt reading (Dietz *et al.*, 2005) as well. Additionally, activation in left parietal and inferior frontal regions has been found to be correlated with reading proficiency (Jobard *et al.*, 2011). The authors attributed the latter effect to phonological processing. Finally, activity in the left parietal lobe has been found to be associated with articulation (e.g. stimulation of the IPS can evoke such intentions; Desmurget *et al.*, 2009), supporting the possibility that activation in the left IPS reflects stronger reliance on phonological processing for written words in SA.

Alternatively, in light of the findings of Chou and colleagues (Chou *et al.*, 2006a, 2009), stronger activation in the left IPS may reflect stronger semantic associations among words in SA. In these studies, activation in the left IPS was found to increase with semantic relatedness (Chou *et al.*, 2006a, 2009). Additionally, when the tasks were presented to children between 9 and 15 years old (Chou *et al.*, 2006b), activation in the left IPS was found to increase with age. This interpretation therefore attributes differences in behavioral measures between LA and SA, as well as differences in activation in the left fusiform, precentral and IFG, to the relative unfamiliarity of written word forms in SA, as did the one presented above. Once words have been successfully decoded, however, increased activation in the left IPS is taken to reflect stronger connections to semantic representations. Such an interpretation would be consistent with words in SA being early acquired and highly familiar, despite the fact that visual presentation of these words is unusual.

The analyses presented demonstrate, therefore, that specific regions, namely the left opercularis and the left fusiform, were more strongly activated in SA when compared with LA, but did not distinguish the activation of SA and Hebrew. This finding suggests that for the unique population of native Arabic speakers, who are both diglossic and bilingual, both the first acquired SA and the later acquired Hebrew at times 'look' like an L2 in the written modality.

Our findings contrast with previous research using auditory stimuli, which indicated that SA words and LA words are processed as L1 and L2, respectively (Ibrahim & Aharon-Peretz, 2005; Ibrahim, 2009). This apparent contradiction stems from the uniqueness of the diglossic situation where language status (as a first or second language) is tightly linked to the modality of presentation. Thus, SA is the first acquired variety, used mainly in spoken language, and therefore occupies a privileged position in processing auditory stimuli. In contrast, LA is acquired later in life, but is then used almost exclusively in the written modality, leading to an advantage in visual word processing, as demonstrated in the current study. Finally, the question remains as to whether the current findings might be better explained in terms of bilingualism and language status, or in terms of effects of familiarity of word forms of the same language across modalities of presentation. Further functional investigations are needed to assess, for example, how low-frequency LA words compare with SA words in the written modality, and how dominance in the auditory vs. visual modality modulates brain activation patterns during SA and LA processing.

Acknowledgements

The authors thank all the participants for their participation in this study. This research was supported by the Israeli Science Foundation (Grant no. 623/11 awarded to A.K. and A.P.) and by the Edmond J. Safra Brain Research Center for the Study of Learning Disabilities. The authors declare that there is no conflict of interest related to the work presented in this paper. This article is dedicated to the memory of Prof. Zvia Breznitz, head of the Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, who passed away on May 18th, 2014.

Abbreviations

AAL, automated anatomical labeling; fMRI, functional magnetic resonance imaging; IFG, inferior frontal gyrus; IPS, intra-parietal sulcus; LA, literary Arabic; LI, laterality index; MFG, middle frontal gyrus; MTG, middle temporal gyrus; ROI, region of interest; RT, reaction time; SA, spoken Arabic; SMA, supplementary motor area; SR, semantically related; SU, semantically unrelated; TE, echo time; TR, repetition time.

References

Abu-Rabia, S. (2000) Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Read. Writ.*, **13**, 147–157.

Abu-Rabia, S. (2001) The role of vowels in reading Semitic scripts: data from Arabic and Hebrew. *Read. Writ.*, **14**, 39–59.

Al-Hamouri, F., Maestú, F., del Río, D., Fernández, S., Campo, P., Almudena, C., García, E., Gonzáles-Marqués, J. & Ortiz, T. (2005) Brain dynamics of Arabic reading: a magnetoencephalography study. *NeuroReport*, **16**, 1861–1864.

Ayari, S. (1996) Diglossia and illiteracy in the Arab world. *Cult. Curriculum*, **9**, 243–253.

Bentin, S. & Ibrahim, R. (1996) New evidence for phonological processing during visual word recognition: the case of Arabic. *J. Exp. Psychol. Learn.*, **22**, 309–323.

Boudelaa, S. & Marslen-Wilson, W.D. (2013) Morphological structure in the Arabic mental lexicon: parallels between standard and dialectal Arabic. *Lang. Cognitive Proc.*, **28**, 1453–1473.

Brett, M., Anton, J.L., Valabregue, R. & Polin, J.B. (2002) Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan. Available on CD-ROM in Neuroimage, 16.

Carreiras, M., Mechelli, A., Estévez, A. & Price, C.J. (2007) Brain activation for lexical decision and reading aloud: two sides of the same coin? *J. Cognitive Neurosci.*, **19**, 433–444.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K. & Miyashita, Y. (2007) Activation of right inferior frontal gyrus during response inhibition across response modalities. *J. Cognitive Neurosci.*, **19**, 69–80.

Chou, T.L., Booth, J.R., Burman, D.D., Bitan, T., Bigio, J.D., Lu, D. & Cone, N.E. (2006a) Developmental changes in the neural correlates of semantic processing. *NeuroImage*, **29**, 1141–1149.

Chou, T.L., Booth, J.R., Bitan, T., Burman, D.D., Bigio, J.D., Cone, N.E., Lu, D. & Cao, F. (2006b) Developmental and skill effects on the neural correlates of semantic processing to visually presented words. *Hum. Brain Mapp.*, **27**, 915–924.

Chou, T.L., Chen, C.W., Wu, M.Y. & Booth, J.R. (2009) The role of the inferior frontal gyrus and inferior parietal lobule in semantic processing of Chinese characters. *Exp. Brain Res.*, **198**, 465–475.

Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C. & Sirigu, A. (2009) Movement intention after parietal cortex stimulation in humans. *Science*, **324**, 811–813.

Dietz, N.A., Jones, K.M., Gareau, L., Zeffiro, T.A. & Eden, G.F. (2005) Phonological decoding involves left posterior fusiform gyrus. *Hum. Brain Mapp.*, **26**, 81–93.

Eviatar, Z. & Ibrahim, R. (2000) Bilingual is as bilingual does: meta-linguistic abilities of Arabic-speaking children. *Appl. Psycholinguist.*, **21**, 451–471.

Ferguson, C. (1959) Diglossia. *Word*, **15**, 325–340.

Field, A. (2005) *Statistics Using SPSS*, 2nd Edn. SAGE Publications Ltd, London.

Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S. & Turner, R. (1995) Analysis of fMRI time-series revisited. *NeuroImage*, **2**, 45–53.

Friston, K., Ashburner, J., Kiebel, S., Nichols, T. & Penny, W. (2007) *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, London.

Gollan, T.H., Forster, K.I. & Frost, R. (1997) Translation priming with different scripts: masked priming with cognates and noncognates in Hebrew-English Bilinguals. *J. Exp. Psychol. Learn.*, **23**, 1122–1139.

Heim, S., Wehnelt, A., Grande, M., Huber, W. & Amunts, K. (2013) Effects of lexicality and word frequency on brain activation in dyslexic readers. *Brain Lang.*, **125**, 194–202.

Ibrahim, R. (2009) The cognitive basis of diglossia in Arabic: evidence from a repetition priming study within and between languages. *Psychol. Res. Behav. Manage.*, **2**, 93–105.

Ibrahim, R. & Aharon-Peretz, J. (2005) Is Literary Arabic a second language for native Arab speakers? Evidence from semantic priming study. *J. Psycholinguist. Res.*, **34**, 51–70.

Jobard, G., Vigneau, M., Simon, G. & Tzourio-Mazoyer, N. (2011) The weight of skill: inter-individual variability of reading related brain activation patterns in fluent readers. *J. Neurolinguist.*, **24**, 113–132.

Joubert, S., Beauregard, M., Walter, N., Bourgouin, P., Beaudoin, P., Leroux, J.M., Karama, S. & Roch Lecours, A. (2004) Neural correlates of lexical and sublexical processes in reading. *Brain Lang.*, **89**, 9–20.

Keatley, C.W., Spinks, J.A. & de Gelder, B. (1994) Asymmetrical cross-language priming effects. *Mem. Cognition*, **22**, 70–84.

Khateb, A., Michel, C.M., Pegna, A.J., O'Dochartaigh, S.D., Landis, T. & Annoni, J.M. (2003) Processing of semantic categorical and associative relations: an ERP mapping study. *Int. J. Psychophysiol.*, **49**, 41–55.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F. & Baker, C.I. (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.*, **12**, 535–540.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W. & Ladurner, G. (2004) The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *NeuroImage*, **21**, 946–953.

Lenartowicz, A., Verbruggen, F., Logan, G.D. & Poldrack, R.A. (2011) Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. *J. Cognitive Neurosci.*, **23**, 3388–3399.

Levin, I., Saiegh-Haddad, E., Hende, N. & Ziv, M. (2008) Early literacy in Arabic: an intervention among Israeli Palestinian kindergartners. *Appl. Psycholinguist.*, **29**, 413–436.

Meschyan, G. & Hernandez, A.E. (2006) Impact of language proficiency and orthographic transparency on bilingual word reading: an fMRI investigation. *NeuroImage*, **29**, 1135–1140.

Rüschemeyer, S.A., Zysset, S. & Friederici, A.D. (2006) Native and non-native reading of sentences: an fMRI experiment. *NeuroImage*, **31**, 354–365.

Saiegh-Haddad, E. (2003) Linguistic distance and initial reading acquisition: the case of Arabic diglossia. *Appl. Psycholinguist.*, **24**, 431–451.

Saiegh-Haddad, E. (2005) Correlates of reading fluency in Arabic: diglossic and orthographic factors. *Read. Writ.*, **18**, 559–582.

Saiegh-Haddad, E., Levin, I., Hende, N. & Ziv, M. (2011) The linguistic affiliation constraint and phoneme recognition in diglossic Arabic. *J. Child Lang.*, **38**, 297–315.

Seghier, M.L., Lazeyras, F., Pegna, A.J., Annoni, J.M., Zimine, I., Mayer, E., Michel, C.M. & Khateb, A. (2004) Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Hum. Brain Mapp.*, **23**, 140–155.

Seghier, M.L., Lazeyras, F., Pegna, A.J., Annoni, J.M. & Khateb, A. (2008) Group analysis and the subject factor in functional magnetic resonance imaging: analysis of fifty right-handed healthy subjects in a semantic language task. *Hum. Brain Mapp.*, **29**, 461–477.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. & Joliot, M. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, **15**, 273–289.

Wartenburger, I., Heekeren, H.R., Abutalebi, J., Cappa, S.F., Villringer, A. & Perani, D. (2003) Early setting of grammatical processing in the bilingual brain. *Neuron*, **37**, 159–170.

Worsley, K.J. & Friston, K.J. (1995) Analysis of fMRI time-series revisited–again. *NeuroImage*, **2**, 173–181.