



Functional reorganization of the reading network in the course of foreign language acquisition

Clara Kuper^{a,b,1}, Jacek Matuszewski^{a,1}, Anna Banaszkiewicz^a, Łukasz Bola^{a,e},
 Michał Szczepanik^a, Małgorzata Draps^{a,c}, Przemysław Kordos^d, Marcin Szwed^e,
 Katarzyna Jednoróg^f, Artur Marchewka^{a,*}

^a Laboratory of Brain Imaging, Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warsaw, Poland

^b Department of Psychology and Educational Sciences, Freie Universität Berlin, Berlin, Germany

^c Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland

^d Department Faculty of Artes Liberales, University of Warsaw, Warsaw, Poland

^e Institute of Psychology, Jagiellonian University, Kraków, Poland

^f Laboratory of Language Neurobiology, Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warsaw, Poland

A B S T R A C T

During foreign language acquisition neural representations of native language and foreign language assimilate. In the reading network, this assimilation leads to a shift from effortful processing to automated reading. Longitudinal studies can track this transition and reveal dynamics that might not become apparent in behavior. Here, we report results from a longitudinal functional magnetic resonance imaging (fMRI) study, which tracked functional changes in the reading network of beginning learners of Greek over one year. We deliberately chose Greek as foreign language that would have similar orthographic transparency but a different alphabet than the native language (Polish). fMRI scans with lexical and semantic decision tasks were performed at five different time points (every ~3 months). Classical language areas (the left inferior frontal gyrus, the left precentral gyrus, and the bilateral supplementary motor cortex), and cognitive control areas (left inferior parietal lobe and bilateral anterior cingulate cortex) showed stronger activation after the first months of instruction as compared to the activation before instruction. This pattern occurred in both tasks. Task-related activity in the reading network remained constant throughout the remaining 6 months of learning and was also present in a follow-up scan 3 months after the end of the course. A similar pattern was demonstrated by the analysis of convergence between foreign and native languages occurring within the first months of learning. Additionally, in the lexical task, the extent of spatial overlap, between foreign and native language in Broca's area increased constantly from the beginning till the end of training. Our findings support the notion that reorganization of language networks is achieved after a relatively short time of foreign language instruction. We also demonstrate that cognitive control areas are recruited in foreign language reading at low proficiency levels. No apparent changes in the foreign or native reading network occur after the initial 3 months of learning. This suggests that task demand might be more important than proficiency in regulating the resources needed for efficient foreign language reading.

1. Introduction

The ability to learn and thereby functionally reorganize the brain is not restricted to critical periods but persists throughout life. Due to the increasing demand to work and communicate in more than one language, foreign language literacy has become a crucial learning achievement, both for integration in a new country and professional success (August and Shanahan, 2006). It remains an open question, how language networks in the brain support reading a foreign language at different learning stages. Among proposed activity patterns for this representation are assimilation and convergence of the native and the foreign language networks. Assimilation refers to changes over time which reduce differences in networks supporting foreign and native language reading. Assimilation is one form of learning-driven functional reorganization in the brain. The term convergence refers to the convergence

hypothesis Green (2003) and means any overlap of task-related neural activity in native and foreign language at one time point.

Across languages, proficient reading in children is characterized by convergence of neural structures used for written and spoken word processing (Chyl et al., 2018; Rueckl et al., 2015). A comparable pattern of speech-text convergence has recently been found for second language (L2) reading (Brice et al., 2019). Brice and colleagues argued that in more proficient readers, computational demands shift to more automated processing. Consequently, first language (L1) and L2 reading networks assimilate with increasing proficiency. While cross-sectional studies as the ones above can pinpoint differences between groups and conditions, only experiments with longitudinal designs can track performance and brain reorganization during new skills acquisition (Draganski et al., 2004; Draganski et al., 2014; Kelly and Garavan, 2005; Li et al., 2014; Siuda-Krzywicka et al., 2016; Zatorre et al., 2012).

Earlier longitudinal studies on reading in L2 investigated mostly functional changes related to L2 proficiency. Stein and colleagues

* Corresponding author.

E-mail address: a.marchewka@nencki.edu.pl (A. Marchewka).

¹ These authors contributed equally.

(2009) demonstrated that higher activation in right and left inferior frontal gyri (IFG) for L2 compared to L1 words was reduced after 5 months of immersive (German) learning and concluded that higher proficiency reduces the need of frontal cognitive control networks. Grant and colleagues (2015) compared the predictions of two models for L2 learning: a developmental extension of the bilingual interactive activation (BIA-d, Grainger et al., 2008) and the convergence hypothesis (Green, 2003). In the BIA-d model, early stages of bilingual word recognition include a direct link between L1 and L2. Later, as the relation between semantics and L2 is reinforced, the connections between L1 and L2 become inhibitory rather than activating to limit cross-language interference. In contrast to this model, the convergence hypothesis predicts that L2 is processed in the same networks as L1 and that similarity between their activity patterns increases as L2 proficiency grows. According to Del Maschio and Abutalebi (2017), early learners recruit additional cognitive control structures along with classical L1 language network. Activity in cognitive control structures decreases with increasing proficiency, resulting in stronger similarity between L1 and L2. Grant and colleagues (2015) examined 24 native English speakers, all classroom-learners of Spanish, and tested them twice with a 4–5-month gap between sessions. Overall, they observed less activation in control areas after the learning period, particularly for unambiguous Spanish words compared to unambiguous English words and concluded that their data was better explained by the convergence hypothesis.

In a more recent longitudinal fMRI study, 12 native English monolingual speakers were enrolled in a highly immersive 12-week intensive French course (Barbeau et al., 2017). In comparison with a control condition (in which participants looked at a string of x's arranged in a sentence-like structure), reading sentences in L1 activated the bilateral IFG and the superior temporal gyri (STG), the insula, precentral regions, and the cerebellum, as well as the inferior parietal lobe (IPL), with stronger activation in the left compared to the right hemisphere. L2 processing invoked similar patterns of activation. Post- vs. pre-training comparison for L2 revealed increased activation in the left IPL and decreased activation in clusters located in the right hemisphere. The authors concluded that activity of the left IPL has a particularly important role in L2 reading, especially due to its positive correlation with reading speed.

Overall, previous longitudinal fMRI studies on L2 acquisition have shown that early L2 processing consistently recruits a more extended language network that mainly diverges from the L1 network by entailing additional cognitive control areas. The observed assimilation of L1 and L2 activities is generally in line with the convergence hypothesis (Abutalebi and Green, 2007; Del Maschio and Abutalebi, 2017; Green, 2003; Green et al., 2006). However, all above-mentioned studies included language pairs with the same alphabet from the family of Latin, Germanic or Roman languages (i.e. German, French, Spanish, Catalan, English). Therefore, reading strategies from L1 could be easily transferred to L2 reading, especially in low-proficient readers (Jacobs and Ziegler, 2015). Similarly, when learning a foreign language, the structure of the native language network is transferred to the foreign language. In Das et al (2011), simultaneous learners of Hindi (transparent) and English (opaque) showed different activity patterns when reading in the respective languages. Specifically, the left inferior gyrus was more activated when reading the opaque language, and the left inferior parietal lobe was more activated for the transparent language. When the languages were learnt sequentially (Hindi as L1, English as L2), the left inferior parietal lobe was activated for both languages. Meschyan and Hernandez (2006) showed similar effects for English - Spanish bilinguals with higher proficiency in their L2 (English) than their L1 (Spanish). The authors suggested that the orthographic network shows less neuroplasticity than other networks involved in language processing like the motor control network. Assimilation of L1 and L2 networks is repeatedly associated with proficiency in longitudinal studies and cross-sectional studies. Meta-analyses show that differences in L2 and L1 pro-

cessing remain even in early bilinguals and suggest that proficiency, rather than age of acquisition is a driving force between L1 and L2 differences (Liu and Cao, 2016)

Along these lines, differences between L1 and L2 reading are also observed when one of the languages was alphabetical (e.g. English) and the other logographic (e.g. Chinese) (Kim et al., 2015; Szwed et al., 2014). In these cases, it has been found that proficient reading is characterized by an accommodation pattern, rather than showing convergence of the foreign language on the native language structures. In the definition by Perfetti and Liu (2005), accommodation refers to a typical native-like activity pattern (e.g. activation of higher-level visual areas) occurring in foreign readers. This shows that foreign language learning is a multidimensional task and should be understood as a complex process, modulated by many factors, amongst others orthographic transparency, linguistic distance, the degree of exposure, learning strategy and age of acquisition (DeLuca et al., 2019; Oliver et al., 2016; Morgan-Short et al., 2012; Ghazi-Saidi and Ansaldo, 2017). To follow this process throughout different stages of acquisition, it is insufficient to compare only two time points.

In the present study we focus on the dynamics of foreign language literacy over a period of 12 months in the reading network supporting the foreign language. The study was conducted with native Polish speakers enrolled in a Greek language course. Five neuroimaging sessions were conducted: one pre-exposure scan, three within-learning-period scans, and one follow-up scan (Fig. 1A). Greek language was chosen because it is written in a different alphabet, hindering an early transfer of native language reading strategies on foreign language but has a comparable orthographic depth to Polish, minimizing differences in native and foreign language processing due to differences in the letter-to-sound mapping. Due to the alphabetic nature and the small difference in orthographic depth between the languages, we can expect an assimilation of the two language networks.

We employed lexical and semantic decision tasks in order to study the relation between native language and foreign language at different levels of linguistic processing: word recognition and reading comprehension. In each task, a reading condition (experimental) is contrasted against a visual search condition (control) revealing reading-related processes, such as grapheme identification and grapheme-phoneme mapping (Coltheart et al., 2001). These processes may also recruit cognitive control regions when needed, such as the prefrontal cortex, the anterior cingulate cortex, the basal ganglia, and the inferior parietal lobe (Abutalebi, Green, 2007; Abutalebi et al., 2008), monitoring conflict between two languages (e.g. Stein et al., 2009). We were interested in a comparison between native and foreign reading networks and to what extent their neural signatures would converge at different stages of learning. Since participants were inexperienced with the Greek alphabet at the beginning of the course, we expected no language-related activity in the pre-exposure scan where written stimuli would not be understood.

We attempted to keep the task demand constant at each time point (apart from the pre-exposure scan) by adapting the stimuli to the vocabulary learned earlier. Thus, even though proficiency would increase over time as participants continue to attend language course, the difficulty of the task performed in the scanner remained comparable. After the first months of instruction (equivalent to low proficiency levels), we expected higher activity in the foreign reading network compared to native reading network including additional activity in cognitive control structures, associated with non-automated reading in the foreign language. During continuation of the language course, we expected an assimilation of native and foreign reading networks, in line with the convergence hypothesis. Drawing from Grant et al. (2015) findings of decreased activity at higher levels of proficiency, we expected this convergence to happen due to decreased activity in the foreign reading network. Overall, we expected activity to change during the time of instruction in foreign, but not native language tasks.

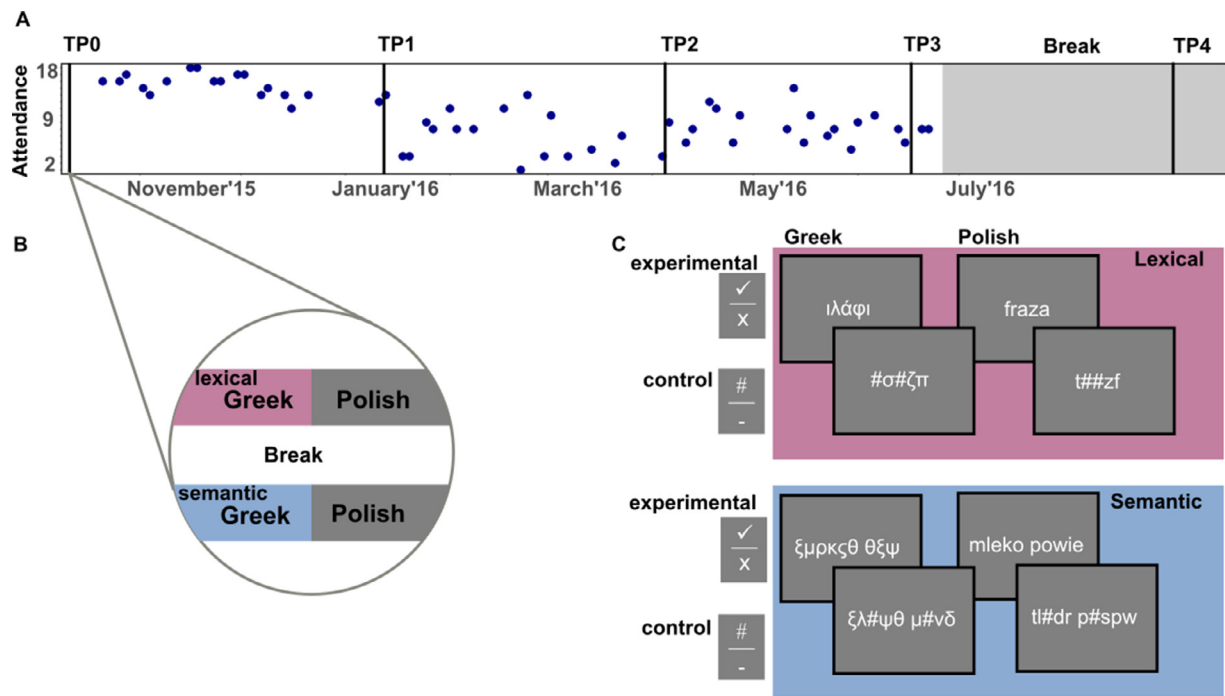


Fig. 1. Overview of Greek language course and fMRI design. **A)** Participants attendance to the Greek course between November and July. The number of participants present at each lecture is plotted on the y-axis. Thick lines denote times when participants were invited for fMRI sessions. **B)** Structure of one fMRI session: participants completed the lexical decision task in both languages in the first run and the semantic judgment task for both languages in the second run. **C)** Example stimuli from the experimental and control conditions in both tasks and languages. Participants were cued with the symbol denoting each condition, left of the colored boxes.

2. Methods

2.1. Participants

Thirty-five women (mean age = 22.1 years, SD = 1.83, range = 20–26 years) were enrolled in the study. Thirteen subjects dropped out due to medical or personal reasons and four were excluded from the analysis due to incomplete data. Therefore, 18 participants were included in the final data analysis (mean age = 22.8 years, SD = 1.97, range = 20–25.8 years).

Subjects were healthy, right-handed, naïve to Greek language, native Polish speakers with normal or corrected-to-normal vision and had no formal musical education. All participants were students of non-linguistic studies and spoke at least one other language (e.g. English, German, French or Russian) in addition to their native language. All foreign languages participants learned before the course came from language families that were unrelated to Greek. Because of the high linguistic distance, we don't assume that any of these languages facilitated the Greek learning process. During the study no other new languages were learned besides Greek. Subjects were screened for MRI safety, gave written informed consent and were paid for participation in the course and for MRI scans. The study was approved by the Committee for Research Ethics of the Institute of Psychology of the Jagiellonian University.

2.2. The Greek course

For the purpose of the experiment we organized an 8-month long Greek course, led in a classroom setting by two professional non-native instructors and, occasionally, a native speaker. The course was mainly based on the workbook “Greek For You Bilingual Greek language books for adolescents & adults” (<https://www.neohel.com/greek-for-you/>) and included alphabet, grammar and a set of vocabulary. At the end, learners reached A1 proficiency level, being able to describe immediate matters and environment, hold a conversation or comprehend a simple monologue. To monitor their learning, six vocabulary class-

room tests (~ every 5 weeks) were performed. Each test consisted of 30 words introduced before the respective test. The classes lasted 1.5 h and took place twice a week. On average, participants attended 33.8 lectures (SD = 3.9), not including homework and personal studies. See Fig. 1A for the time course of the Greek language course.

2.3. Study design and fMRI tasks

In total, participants were invited for five fMRI sessions - time points (TP). One session took place at pre-exposure (TP0), three sessions during the course (TP1 to TP3) and a follow-up session (TP4), 3 months after the end of the course. On average, the total time of instruction was 0 h at TP0, 24.75 h (SD = 0.23 h) at TP1, 38.58 h (SD = 0.37 h) at TP2, and 51.1 h (SD = 0.34 h) at TP3 and TP4. Scans were performed within approximately 3-months long intervals. fMRI tasks were based on two types of language processing: lexical and semantic, presented in two conditions – experimental and control (Fig. 1B and 1C), in both Greek (foreign) and Polish (native). In the experimental condition of the lexical decision task subjects had to discriminate words from pseudowords. Pseudowords were created by changing one letter of an existing word. The semantic decision task was conducted in a similar manner with semantically correct or deviant two-words sentences, and the subjects having to indicate their correctness. As an example, the sentence “mum speaks” would be semantically correct, while the sentence “table speaks” is semantically incorrect. In the control conditions to both tasks, random consonant strings length-matched to the task stimuli (i.e. one string for lexical and two strings for semantic decision tasks) were displayed on the screen. Half of the strings contained two “#” symbols – participants had to report whether they were present among letters (see Fig. 1, and Table 1 for stimuli examples).

2.4. Procedure

Tasks were presented within 2 separate runs (one for each task) with 4 language (foreign / native) x condition (experimental / control) com-

Table 1
Example of text stimuli presented in the fMRI procedure.

Stimulus lexical	Translation	Language	Condition	Stimulus semantic	Translation	Language	Condition	Correct
κράνος	helmet	Greek	Exp	τάντα ακούει	bag hears	Greek	Exp	Incorrect
Νιάτα	youth	Greek	Exp	θεία νιώθει	aunt feels	Greek	Exp	Correct
Σέλει	pseudoword	Greek	Exp	πλοίο ράβει	ship stitches	Greek	Exp	Incorrect
Δέκα	ten	Greek	Exp	κύριοι είδαν	(the) gentlemen saw	Greek	Exp	Correct
σκήθος	pseudoword	Greek	Exp	αλεπού είπε	fox said	Greek	Exp	Incorrect
Urlop	Vacation	Polish	Exp	idzie las	forest walks	Polish	Exp	Incorrect
Rower	Bicycle	Polish	Exp	witać ludzi	greet people	Polish	Exp	Correct
Drzwu	pseudoword	Polish	Exp	znają ciocię	(they) know (the) aunt	Polish	Exp	Correct
chepia	pseudoword	Polish	Exp	banan gra	banana plays	Polish	Exp	Incorrect
Onej	pseudoword	Polish	Exp	test kocha	(the) test loves	Polish	Exp	Incorrect
#κρ#θβ	-	Greek	Con	σψ λςνφ ξφπξ	-	Greek	Con	-
Βρχτβ	-	Greek	Con	τσψντψ λξφπξφ	-	Greek	Con	-
βπθ#ρ#	-	Greek	Con	πφνξφ #ψνφ#	-	Greek	Con	-
Zκβνψ	-	Greek	Con	ψλξπ#ζ ξ#δξ	-	Greek	Con	-
#ζ#πς	-	Greek	Con	σκζλ## βρθκξ	-	Greek	Con	-
l#m#z	-	Polish	Con	#hám#k jzł	-	Polish	Con	-
ćwbnm	-	Polish	Con	bódz sńák	-	Polish	Con	-
Ćlśz	-	Polish	Con	złłz grłzśz	-	Polish	Con	-
ń#	-	Polish	Con	młz ##zł	-	Polish	Con	-
żzhápw	-	Polish	Con	mrhczł pśznf	-	Polish	Con	-

binations per run. In each run, Greek was presented first, followed by Polish. Stimuli were presented in a blocked design. Within each language, five blocks of each condition were alternated. The order of blocks (experimental or control first) was counterbalanced between participants. Each block consisted of 8 trials which lasted 4 s in lexical decision task, and 5 s in semantic decision task. Each block started with a condition cue presented for 1, 1.5 or 2 s followed by a blank screen for 6–8 s. In each trial, a fixation cross was shown for 1 s at the beginning of each trial and stimuli were presented for 1 s (2 s in semantic decision task) afterwards. A blank screen was presented for the remaining 2 s. Participants could answer as soon as the stimuli came on screen, and until the trial ended, but were never directly instructed to give answers as quickly as possible. Responses not given before the end of the trial were counted as misses. Within one run, 160 trials were completed in total, 80 trials for each language, and 40 trials for each condition. The total duration of the lexical decision task was 13 min (block duration = 32.1 s.; stimuli length = 1 s.; answer window = 2 s; inter-stimulus interval = 1 s.) and of the semantic decision task 15.7 min (block duration = 40.1 s.; stimuli length = 2 s.; answer window = 2 s; inter-stimulus interval = 1 s.).

2.5. Stimuli

All words and pseudowords were 3–6 letters long ($M = 4.9$ letters, $SD = 0.94$). Polish and Greek stimuli were controlled for frequency, and orthographic neighborhood according to the SUBTLEX-PL database (Mandera et al., 2015). In the semantic decision task, grammatically valid sentences with the construction Verb + Noun or Noun + Verb were used. The set of Greek stimuli was based on the list of Polish stimuli and the workbook used during course instructions.

During the learning period, difficulty-adapted Greek stimuli sets were presented, such that the cognitive demand, reflected by participant's performance, would be comparable across time points (see Table 1 for examples). Therefore, for each period between fMRI sessions a set of stimuli was selected and taught in class, resulting in 4 different, thematically structured sets of stimuli. In each fMRI session, the presented stimuli were randomly chosen from all sets learned up to the respective session (from set 1 at TP1, from set 1 and 2 at TP2, etc.). At TP0, vocabulary from all sets was included, all unknown at this time point. No word occurred twice in subsequent time points in the lexical decision task, some words occurred twice in the semantic decision task, due to the limited set of available stimuli. Stimuli were displayed using Presentation software (Neurobehavioral Systems, Berkeley, CA), in the center of an MRI compatible screen with grey background.

2.6. Imaging parameters

Data was collected on a 3T Trio Trim (Siemens) MRI Scanner using a 12-channel head coil. T1-weighted images were acquired with the following parameters: 176 slices, slice-thickness = 1 mm, TR = 2530 ms, TE = 3.32 ms, flip angle = 7°, FOV = 256 mm, matrix size: 256 × 256, voxel size: 1 × 1 × 1 mm. For functional imaging, 41 slices with a slice-thickness of 3 mm, TR of 2.5 s, flip angle of 80°, FOV of 216 × 216 mm and a voxel size of 3 × 3 × 3 mm were collected in each volume.

2.7. Data preprocessing and analyses

Behavioral results were analyzed in R (version 3.4.2, The R Foundation for Statistical Computing) using R studio (Version 1.1.383, RStudio, Inc.). We used repeated-measure ANOVA (rmANOVA) with condition (experimental vs. control), language (foreign vs. native), task (semantic vs. lexical), and time (factorial TP0 – TP4), as well as their interactions as independent variables and performance as a dependent variable. Subjects were also included as a grouping factor. Data was not normally distributed (Wilcoxon Rank-Signed Test, $V=259560$, $p < 0.05$) and variance was inhomogeneous for the factors language (Levene-Test, $Df = [1,718]$, $F = 613.79$, $p < 0.05$) and condition (Levene-Test, $Df = [1,718]$, $F = 592.57$, $p < 0.05$). Thus, additional non-parametric Sheirer-Hare Tests were conducted (Holmes et al., 2016). For interpretability and comparability, we report the results of the rmANOVA, given that the same effect was observed from the Sheirer-Hare Test. Wilcoxon rank-signed tests were used to compare performance between time points, and V-values are reported from that statistic. One-sample t-tests are used to test if a given condition deviates from chance level.

The pre-processing and statistical fMRI analyses were performed using SPM12, version 6906 (Wellcome Imaging Department, University College, London, UK, <http://fil.ion.ucl.ac.uk/spm>), run in MATLAB R2017b (The Math-Works Inc. Natick, MA, USA). All functional images were realigned, and motion corrected. Analysis of head motion was performed using the NITRIC artifact detection toolbox (https://www.nitrc.org/projects/artifact_detect/). Motion of more than 5 mm in 4 or less volumes over all sessions was detected in 4 participants. Volumes were not excluded due to motion but realigned together with the other volumes to account for shifts in the image. A structural longitudinal registration toolbox (Ashburner and Ridgway, 2013) was used to implement an averaged T1-weighted (T1-w) image from all scans. Further on, functional images were normalized to MNI (Montreal Neurological Institute) space using deformation fields of the averaged

T1-w images. Functional images were smoothed with a 6 mm full width at half maximum Gaussian kernel.

At the subject level, each condition (experimental, control) and language (foreign, native) was modeled with an independent General Linear Model (GLM). Timings of each condition as well as instruction screens between blocks were modeled with boxcar functions, starting with the first presentation of a fixation cross. Additional six head movement regressors were included in the design matrix. All boxcar functions were convolved with the hemodynamic response function (HRF) as implemented in SPM12. High-pass filter of 128 Hz was applied to remove low-frequency drift from the imaging data. Next, we computed *experimental* > *control* contrasts at the single subjects level for each language and time point. These contrasts were entered in all models at the second level.

Our analysis was conducted using two approaches. First, we tested how functional activity levels depend on our conditions: language and time in each of the tasks. Thereby, we answer the question if the activity in a given region changes with longer learning periods. Next, we tested, specifically for spatial convergence, in order to see if the location of activity in the brain changes with learning of a foreign language.

2.8. Functional activity analysis

We tested our hypothesis that activity in response to the tasks would change in foreign reading, but not in native reading over time, by computing interaction effects in a within-subject ANOVA with language (2) and time (5) factors. For further analysis a mask from this contrast was applied. We explored the interaction effects for each language separately in a within-subject ANOVA with language (1) and time (5) factors. Comparison between time points - and thus, the effect of one 3-month training block - was performed with paired t-tests.

Task-related responses were considered significant at a voxel-wise threshold $p < 0.001$, corrected for multiple comparisons across the whole brain using a cluster-level Family Wise Error (FWE) at threshold $p < 0.05$. For anatomical identification the *xjview* toolbox (<http://www.alivelearn.net/xjview>) was used, based on the Automatic Anatomic Labeling plug-in for SPM (Tzourio-Mazoyer et al., 2002).

To visualize the pattern of signal differences across time, region of interest (ROI) analyses were built of the clusters showing higher activity in TP1 compared to TP0 in the foreign language. Next, the voxel-time course was extracted from these clusters in the 2×5 ANOVA language $\times$ time. This step was done for visualization purposes only and no statistical analysis was performed on the time course of these regions. The extraction of the voxel time-course was done in MarsBaR toolbox (Brett et al., 2000).

Finally, additional analyses were performed on four independent ROIs. Three regions were defined anatomically based on the automated anatomical labeling (AAL) atlas. The tested regions were: left Broca's area (consisting of combined masks from pars triangularis and opercularis of left inferior frontal gyrus) and bilateral inferior parietal lobules. Lastly, the lexical-sensitive part of the visual word form area ROI was created, based on Lerma-Usabiaga et al., 2018 (defined as an 8 mm radius sphere at MNI coordinates [-42, -58, -10]). The ROIs were chosen with respect to the following criteria: visual word form area because of its relevance in visual language processing (Dehaene (2010)), Broca's area because of its general importance in language processing and foreign language processing (Stein et al., 2009), left inferior parietal lobule because of its assumed role in a cognitive control network (Abutalebi and Green, 2007) and its importance in the study by Barbeau et al (2017) and finally the anterior part of cingulate cortex, as a further central part of the cognitive control network (Abutalebi et al., 2008).

To capture the full effect of learning in task-relevant regions, we computed a 2×2 rmANOVA including language (native and foreign) and time (TP1 and TP3) as factors. These time points were chosen to cover the longest possible learning period after familiarizing with the foreign alphabet (6 months).

2.9. Spatial convergence analysis

In order to visualize the native and foreign language processing networks and explore the spatial convergence on the whole-brain level we performed conjunction analyses. Specifically, for native language, *experimental* > *control* contrasts were averaged across all time points at the subject level and summarized in one sample T tests for lexical and semantic decision tasks respectively. Corresponding foreign language activations were derived from TP0, TP1 and TP3, which allowed us to additionally explore the convergence hypothesis by the means of logical AND conjunction. All maps were corrected for multiple comparisons with Family Wise Error at the cluster level (FWE) of $p < 0.05$ (with initial voxel-wise threshold of $p < 0.001$).

Next, the time course of spatial convergence between foreign and native languages was tested in ROIs described above via Dice coefficients analysis. Dice coefficient equals twice the number of voxels common for both languages divided by the sum of number of voxels activated in each language (Bennet and Miller, 2010). First, for each subject in each TP and task *experimental* > *control* contrasts for foreign and native language activations as well as their conjunctions were thresholded with $p < 0.05$. Next, number of voxels active in each contrast was extracted via *spm_summarise* function from all ROIs. Finally, in order to quantitatively test overlap changes over time, Dice coefficients were calculated and entered into ROI-and-task-specific rmANOVAs with time (5) and subject factors.

Region time courses and ROI analysis were visualized using GraphPad Prism 8.3 (GraphPad Software, La Jolla California USA, www.graphpad.com). Statistical maps from SPM were overlaid on anatomical maps using NeuroElf v 1.1. (build 17080412, (c) Jochen Weber). Final figures were assembled in Inkscape 0.92 (www.inkscape.org).

2.10. Data/code availability statement

Thresholded maps for each 2nd level contrast and behavioral performance are available for downloads: https://lobi.nencki.gov.pl/a/2-greek_supplementary

3. Results

3.1. Behavioral performance analysis

Performance was measured as % correct answers. We found significant main effects of condition ($F(1,17)=725.6$, $p < 0.001$) and language ($F(1,17)=1391$, $p < 0.001$) on performance as well as an interaction between these two factors ($F(1,17)=770.8$, $p < 0.001$). Also, we observed an interaction of task $\times$ condition ($F(1,17)=5.146$, $p < 0.05$) and task $\times$ language ($F(1,17) = 6.317$, $p < 0.05$). There was no main effect of time ($F(1,17)=0.0011$, $p > 0.05$) (Fig. 2).

We had a strong assumption that in the experimental condition of the foreign language participants would perform at chance level at TP0 and above at later time points. Thus, performance in TP0 was compared to performance at other time points. At each given time point, participants performed better than in TP0 (TP1: $V = 0$, $p < 0.001$; TP2: $V = 2$, $p < 0.001$; TP3: $V = 21$, $p < 0.001$; TP4: $V = 30.5$, $p < 0.005$). Performance at TP0 was not different from chance level (0.5) ($Df = 17$, $t = -0.59$, $p > 0.05$). We found the same results for the semantic decision task, with performance at chance level at TP0 ($Df = 17$, $t = 0.07$, $p > 0.05$) and higher performance at TP1 ($V = 0$, $p < 0.001$) and later other time points (TP2: $V = 24$ $p < 0.05$; TP3: $V = 12$, $p < 0.005$; TP4: $V = 30.5$ $p < 0.05$). In both tasks, performance was highest at TP1. For the lexical decision task, mean performance was 49% at TP0 ($SD = 0.09$), 75% at TP1 ($SD = 0.11$), 65% at TP2 ($SD = 0.08$), 60% at TP3 ($SD = 0.07$) and 55% at TP4 ($SD = 0.09$). In the semantic decision task, mean performance was 50% at TP0 ($SD = 0.08$), 70% at TP1 ($SD = 0.12$), 65% at

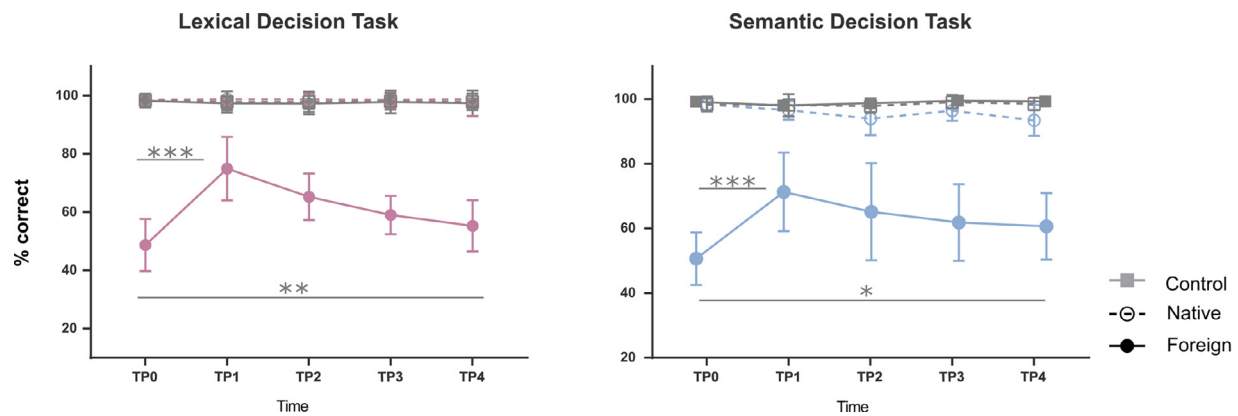


Fig. 2. Behavioral scores in the MRI scanner during both tasks. Note that the difficulty of the behavioral task in the MR scanner was designed to be increasingly difficult at each session; performance is therefore not a direct indicator of proficiency levels. For Polish language (native) and in the control condition (visual search) performance was at ceiling level. Only in the experimental condition in Greek (foreign) performance was modulated by time during the lexical and semantic decision task. Experimental conditions are depicted in violet (lexical) and blue (semantic). Asterisks symbolize a significant change (** $p < 0.001$, ** $p < 0.005$, * $p < 0.05$) between time points as calculated with post-hoc paired t-tests.

TP2 (SD=0.15), 61% at TP3 (SD = 0.12) and 60% at TP4 (SD=0.10). Fig. 2 summarizes the results from the behavioral analysis.

Participants were not instructed to react fast or at a specific time. In both tasks reaction times were shorter for in the experimental condition of the foreign language at TP0 compared to the remaining time points. Details about the reaction time distribution and statistics are reported in the supplementary materials (S1.1, Tables S1 and S2).

3.2. fMRI Results

3.2.1. Functional activity changes

In line with our general hypothesis that functional activity by means of the signal response amplitude changes with time in the foreign, but not in the native language, we found language $\times$ time interaction effects in both tasks. For the lexical decision task, the interaction between language and time was significant in the precuneus (Pre), the frontal pole (including anterior cingulate gyrus), the superior frontal gyrus, and the supplementary motor cortex, bilaterally. In the left hemisphere, the language $\times$ time interaction contrast included clusters in the angular gyrus, the precentral gyrus (expanding to the inferior frontal gyrus), the superior parietal lobule, and the inferior temporal gyrus. In the right hemisphere, clusters were in the lateral occipital cortex and the middle temporal gyrus (Fig. 3A).

The interaction between language and time in semantic decision task involved brain regions in inferior frontal and parietal areas such as the precentral gyrus and the inferior frontal gyrus, the inferior superior parietal lobe, the angular gyrus, the inferior temporal gyrus including the fusiform gyrus, and the lateral occipital cortex in the left hemisphere (Fig. 3D).

To confirm that the language $\times$ time interaction is due to changes specific to the foreign reading, we assessed its directionality in two independent 1×5 ANOVAs, containing only one language each. For the foreign language, main effects of time were observed in both tasks and overlapped with the interaction effect. Supplementary figures in section S1.2 show these overlaps for both tasks. We observed no main effect of time in either task (lexical and semantic) conducted in the native language.

We expected the recruitment of areas associated with the reading network and additional cognitive control structures at early stages of learning. Direct comparison between time points in the foreign language allowed us to describe the time course of activity in this network. This approach also enabled evaluation of the directionality of the activation (activation vs. deactivation). In the case of the lexical decision task, ac-

tivity in the bilateral angular gyrus, the bilateral precuneus cortex, the left superior frontal gyrus and the right frontal pole was stronger at TP0 than at TP1. The inverse contrast, TP1 > TP0, revealed the activation in clusters mainly located in the left hemisphere: in the left precentral gyrus extending to the inferior frontal gyrus, in the left lateral occipital cortex (including superior and inferior parietal lobe and middle and superior occipital lobes), and in the bilateral supplementary motor area (Fig. 3B and Table 2). Pairwise comparison of later time points (TP1 vs TP2, TP2 vs TP3, and TP3 vs TP4) showed no significant effects, with the exception of one cluster in the left middle temporal gyrus that decreased activity between TP3 and TP4 (the follow-up session).

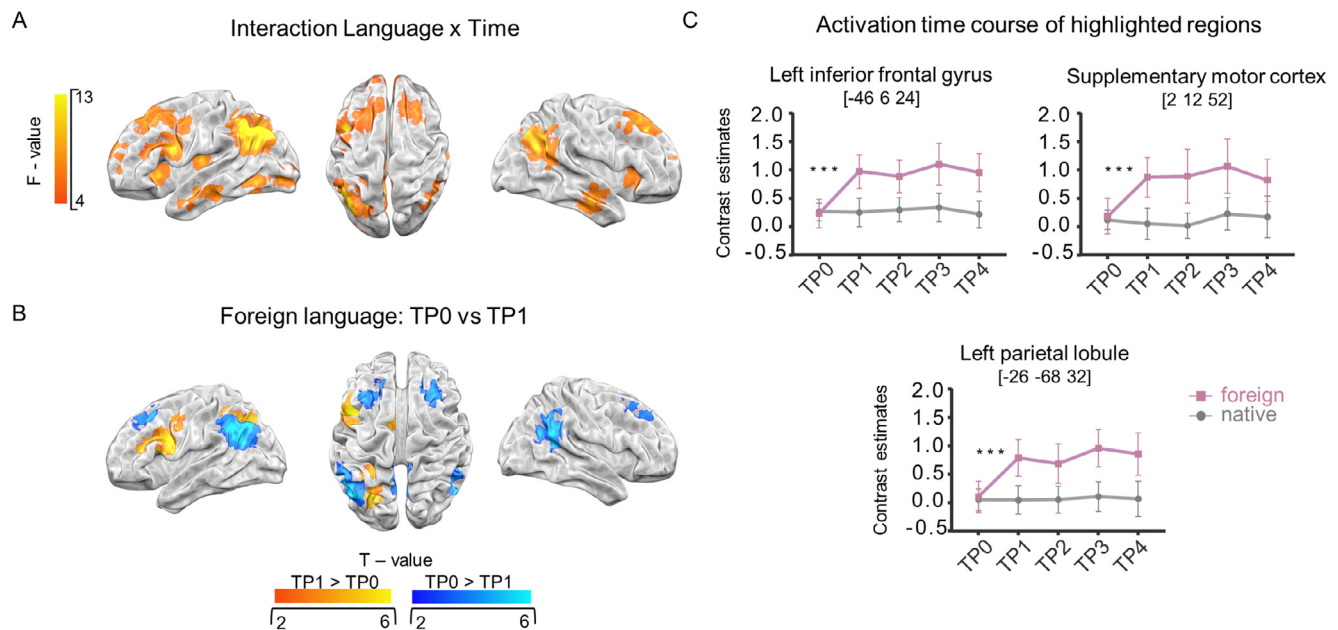
In the case of the semantic decision task - comparable to the lexical decision task, higher activation at TP0 compared with TP1 was observed bilaterally in the angular gyrus. Higher activity at TP1 was present in clusters in the left hemisphere, such as the left insular cortex (including the precentral gyrus and the inferior frontal gyrus), the left inferior and the superior divisions of the parietal lobe (superior and inferior), in the left occipito-temporal fusiform cortex, the bilateral supplementary motor area and the right lateral occipital cortex (Fig. 3E and Table 3). Pairwise comparison of later time points showed no differences between them in either direction.

Regions in which activity changes within the first three months were common to both tasks included the inferior frontal gyrus, the supplementary motor cortex and the left parietal lobe. In all regions, the activity changes over the full experiment followed a characteristic time course, depicted in Fig. 3C and 3F. Activity in all regions associated with a learning related change increased between TP0 and TP1 in foreign language tasks and remained elevated above native language level for the remaining time points.

Furthermore, in order to quantitatively test the stability of brain responses to Greek after TP1, we specified similar ANOVA designs with exclusion of the data from TP0. Results revealed no significant changes in neither lexical or semantic decision tasks.

In our first step, we checked if the amplitude of functional activity assimilates between native and foreign language over time. This assimilation was not observed in the interaction between language and time. To improve detection of changes over time from a conservative whole-brain analysis, we conducted an ROI analysis including central structures of the reading (left Broca's area, the left visual word form area) and the cognitive control (the left inferior parietal lobule and anterior cingulate cortex) networks. We looked for interactions in a 2×2 design, including 2 time points (TP1 and TP3) and two languages (native and foreign). All interactions were not significant.

Lexical Decision Task



Semantic Decision Task

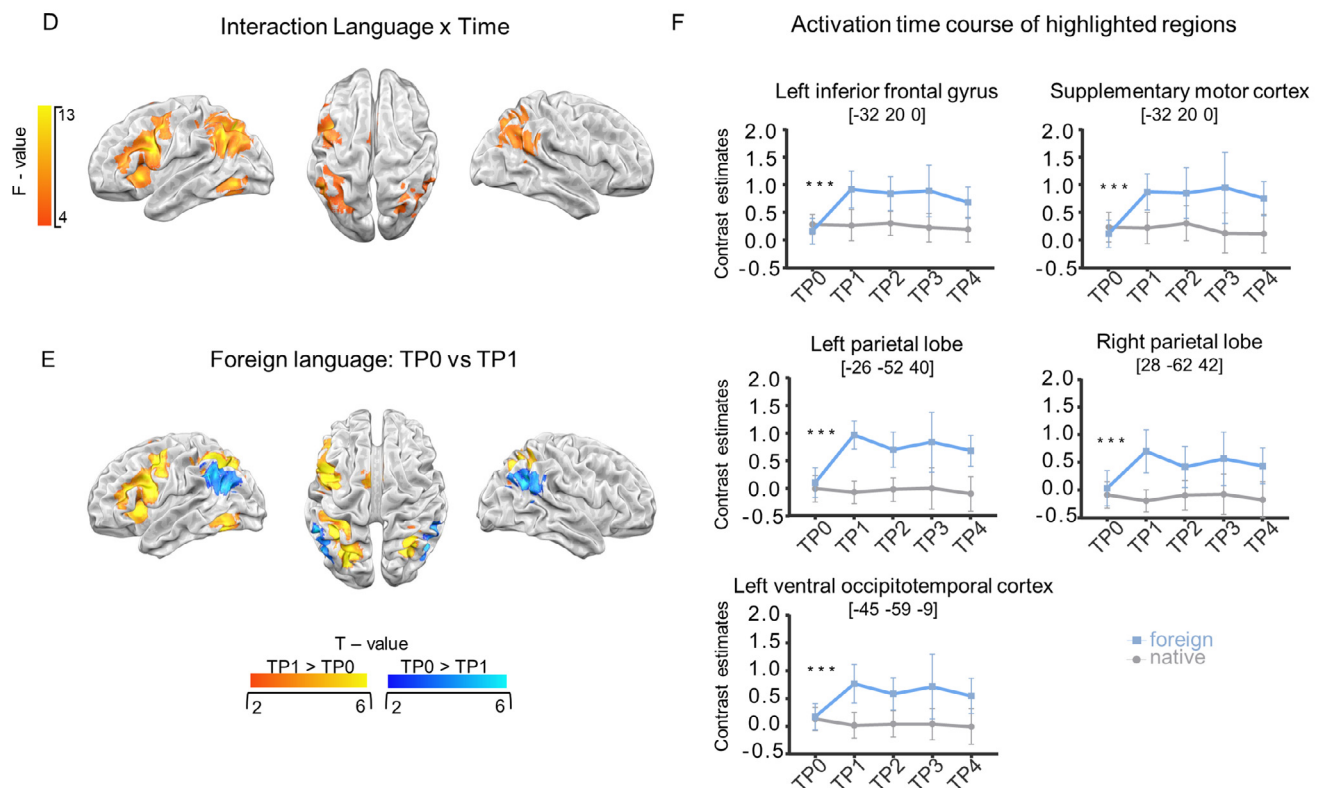


Fig. 3. Activity changes in the foreign language network. Upper panel (A - C) shows results from the lexical decision task, lower panel (D - F) shows results from the semantic decision task. All results represent *experimental* > *control* contrasts in their respective tasks and languages. A) Language x time interaction effects of language and time in the lexical decision task. The network activated in the interaction largely overlaps with the main effect of time in the foreign language (see S1.2). The interaction also involves regions from the default mode network, such as the bilateral angular gyrus. B) Language x time interaction effects can be observed in the contrasts between time points. Differences in activation between TP1 and TP0 are shown. Hot colour maps indicate TP1 > TP0 contrast. Cold colour maps indicate TP0 > TP1 contrast. C) Visualization of the language x time interaction in regions from the paired t-test between TP0 and TP1. D) Interaction effects of Language and Time in the semantic decision task. E) Directional effects between TP0 and TP1. Hot colour maps indicate TP1 > TP0 contrast. Cold colour maps indicate TP0 > TP1 contrast. F) Visualization of the language x time interaction in regions from the paired t-test between TP0 and TP1. Contrast estimates are computed based on the whole cluster with the peak voxels mentioned in the heading. Asterisks (***) symbolize a significant change ($p < 0.001$) between time points as calculated with post-hoc paired t-tests.

Table 2

Results of the lexical decision task. Structures in bold indicate the location of the peak voxel in each cluster. Structures with 20 or more voxels within the same cluster are listed below. For brevity, only clusters with at least 250 voxels are included in the table. The full report can be found in the supplementary materials, section S1.3.

Interaction time x language						
Region Label	Extent	F-value	MNI Coordinates			
			x	Y	z	
Bilateral						
Precuneus Cortex	2290	11.231	-2	-44	42	
Precuneus R, Precuneus L, Mid Cingulate Gyrus L, Mid Cingulate Gyrus R,						
Frontal Pole	1174	8.325	-20	56	20	
Anterior Cingulate Gyrus L, Superior Frontal Gyrus L, Superior Frontal Gyrus R,						
Anterior Cingulate Gyrus R, Superior Frontal Gyrus L						
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)	297	9.884	-4	0	62	
Supplementary Motor Cortex L, Supplementary Motor Cortex R, Middle Cingulate Gyrus R						
Left Hemisphere						
Angular Gyrus	1525	16.351	-54	-54	38	
Angular Gyrus L, Inferior Parietal Lobe L, Middle temporal gyrus L, Middle occipital Lobe L, Supramarginal Gyrus L						
Superior Frontal Gyrus	1010	9.200	-26	28	56	
Middle Frontal Gyrus L, Superior Frontal Gyrus L						
Precentral Gyrus	718	7.581	-58	10	26	
Precentral Gyrus L, Inferior Frontal Gyrus L, pars opercularis, Inferior Frontal Gyrus L, pars triangularis						
Superior Parietal Lobule	503	8.662	-28	-48	40	
Superior Parietal Lobe L, Middle Occipital Cortex L, Inferior Parietal Lobe L, Superior Occipital Cortex						
Inferior Temporal Gyrus, temporooccipital part	265	11.351	-48	-60	-12	
Inferior Temporal Gyrus L, Inferior Occipital Cortex L, Middle Temporal Gyrus L						
Right Hemisphere						
Lateral Occipital Cortex, superior division	1030	9.731	42	-70	40	
Angular Gyrus R, Middle Temporal Gyrus R, Superior Temporal Gyrus R, Inferior Parietal Lobe R						
Superior Frontal Gyrus	754	11.006	22	34	40	
Superior Frontal Gyrus R, Middle Frontal Gyrus R						
Right Cerebral White Matter	335	6.985	52	-18	-26	
Middle Temporal Gyrus R, Inferior Temporal Gyrus R, Superior Temporal Gyrus R						
TP0>TP1						
Region Label	Extent	t-value	MNI Coordinates			
			x	y	z	
Bilateral						
Precuneus Cortex	1091	6.440	8	-56	26	
Precuneus R, Precuneus L, Middle Cingulate Gyrus L, Middle Cingulate Gyrus R, Posterior Cingulate Gyrus R						
Left Hemisphere						
Angular Gyrus	1142	8.137	-46	-58	30	
Angular Gyrus L, Inferior Parietal Lobe L, Middle Temporal Gyrus L, Supramarginal Gyrus L						
Superior Frontal Gyrus	250	4.831	-24	32	42	
Middle Frontal Gyrus L						
Right Hemisphere						
Angular Gyrus	659	7.253	52	-56	20	
Angular Gyrus R, Middle Temporal Gyrus R, Superior Temporal R, Inferior Parietal Lobe R						
TP1 > TP0						
Region Label	Extent	t-value	MNI Coordinates			
			x	y	z	
Bilateral						
Paracingulate Gyrus	284	6.995	2	12	52	
Supplementary Motor Area L, Supplementary Motor Area R						
Left Hemisphere						
Precentral Gyrus	659	8.285	-46	6	24	
Precentral Gyrus L, Inferior Frontal Gyrus, pars opercularis L, Inferior Frontal Gyrus, pars triangularis L						
Lateral Occipital Cortex, superior division	468	7.095	-26	-68	32	
Superior Parietal Lobe L, Middle Occipital Lobe L, Inferior Parietal Lobe L, Superior Occipital Lobe L						

Table 3

Results of the semantic decision task. Structures in bold indicate the location of the peak voxel in each cluster. Structures with 20 or more voxels within the same cluster are listed below. For brevity, only clusters with at least 250 voxels are included in the table. The full report can be found in the supplementary materials, section S1.3.

Interaction time x language						
Region Label	Extent	F-value	MNI Coordinates			
			x	y	z	
Bilateral						
Paracingulate Gyrus	363	11.128	-2	10	52	
Supplementary Motor Area L, Supplementary Motor Area R						
Left Hemisphere						
Precentral Gyrus	2054	16.342	-42	6	28	
Precentral Gyrus L, Inferior Frontal Gyrus, pars triangularis L, Inferior Frontal Gyrus, pars opercularis L, Insula L, Middle Frontal Gyrus L, Rolandic Operculum L						
Left Cerebral White Matter	1267	16.221	-26	-50	40	
Inferior Parietal Lobe L, Superior Parietal Lobe L, Middle Occipital Lobe L, Superior Occipital Lobe L						
Angular Gyrus	711	13.697	-56	-54	38	
Angular Gyrus L, Inferior Parietal Lobe L, Supramarginal Gyrus L, Middle Temporal Gyrus L, Middle Occipital Lobe L						
Inferior Temporal Gyrus, temporooccipital part	425	10.611	-48	-60	-10	
Inferior Temporal Gyrus L, Inferior Occipital Cortex L, Fusiform Gyrus L, Middle Temporal Gyrus L						
Lateral Occipital Cortex, superior division	289	8.392	30	-58	50	
Superior Occipital Cortex L, Inferior Parietal Lobe L, Angular Gyrus L, Superior Parietal Lobe L, Middle Occipital Lobe L						
Right Hemisphere						
Angular Gyrus	434	6.780	52	-56	28	
Angular Gyrus R, Supramarginal Gyrus R, Superior Temporal R						
TP0 > TP1						
Region Label	Extent	t-value	MNI Coordinates			
			x	y	z	
Left Hemisphere						
Supramarginal Gyrus, posterior division	406	6.171	-56	-50	34	
Angular Gyrus L, Inferior Parietal Lobe L, Supramarginal Gyrus L, Middle Temporal Gyrus L						
Right Hemisphere						
Angular Gyrus	321	7.171	48	-54	28	
Angular Gyrus R, Supramarginal Gyrus R, Superior Temporal Gyrus R						
TP1 > TP0						
Region Label	Extent	t-value	MNI Coordinates			
			x	y	z	
Bilateral						
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)	361	11.363	-2	8	54	
Supplementary Motor Area L, Supplementary Motor Area R						
Left Hemisphere						
Insular Cortex	1882	10.756	-32	20	0	
Precentral Gyrus L, Inferior Frontal Gyrus, pars triangularis L, Inferior Frontal Gyrus, pars opercularis L, Insula L, Middle Frontal Gyrus L						
Left Cerebral White Matter	1235	15.903	-26	-52	40	
Inferior Parietal Lobe L, Superior Parietal Lobe L, Middle Occipital Cortex L, Superior Occipital Cortex L						
Temporal Occipital Fusiform Cortex	400	8.839	-44	-62	-14	
Inferior Temporal Gyrus L, Inferior Occipital Gyrus L, Fusiform Gyrus L, Middle Temporal Gyrus						
Right Hemisphere						
Lateral Occipital Cortex, superior division	282	8.291	28	-62	42	
Superior Occipital Cortex R, Inferior Parietal Lobe R, Angular Gyrus R, Superior Parietal Lobe R, Middle Occipital Cortex R						

3.2.2. Spatial convergence

Whole brain conjunction analysis for the lexical task indicated significant foreign - native overlap in the bilateral supplementary motor area at TP0 (Fig. 5, upper panel). Additionally, small overlap was found in the left IFG (pars triangularis). At TP1 overlaps additionally included bilateral insular cortex, bilateral inferior frontal gyri, left inferior parietal cortex, left ventral occipitotemporal cortex and left pallidum. These patterns remained mostly stable at TP3, indicating no major changes in

spatial convergence between these two TPs (Fig. 5 left, middle and lower panels). Similar pattern was observed for the semantic task. Specifically, common activations between native and foreign languages at TP0 included the bilateral supplementary motor area. Then, at TP1 conjunctions were observed also in the bilateral early visual and insular cortex, the left inferior frontal gyrus, the left inferior temporal gyrus, the left precentral gyrus and in subcortical structures including the left pallidum and the thalamus. Again, a similar pattern was observed at TP3 (Fig. 5

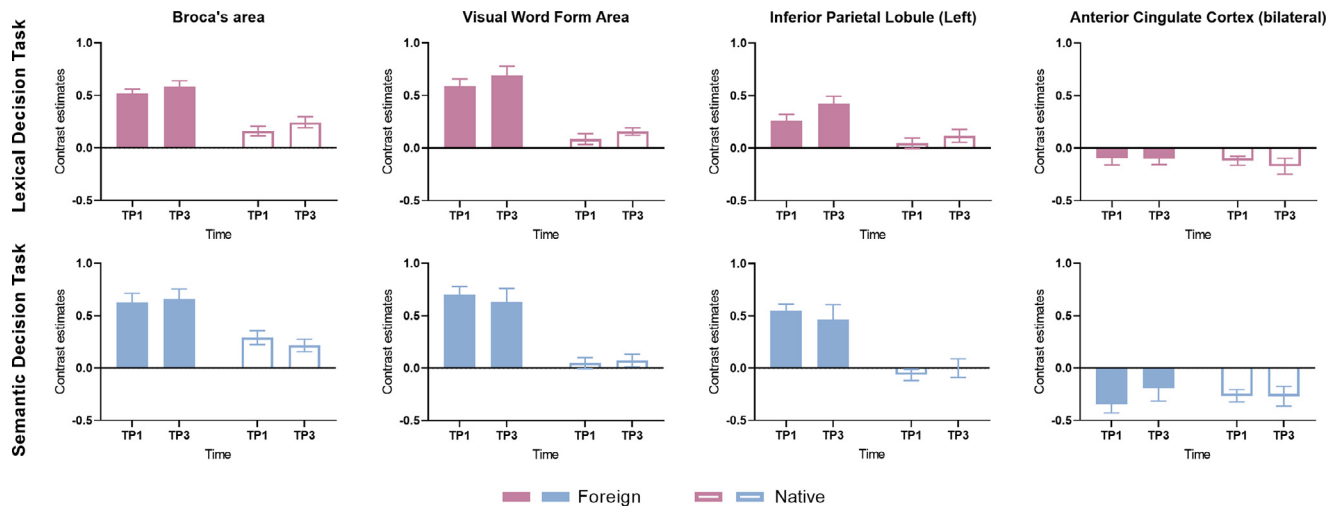


Fig. 4. Effect of time and language in selected Regions of Interest. TP = time point. Values represent *experimental* > *control* contrasts (see Method section for details).

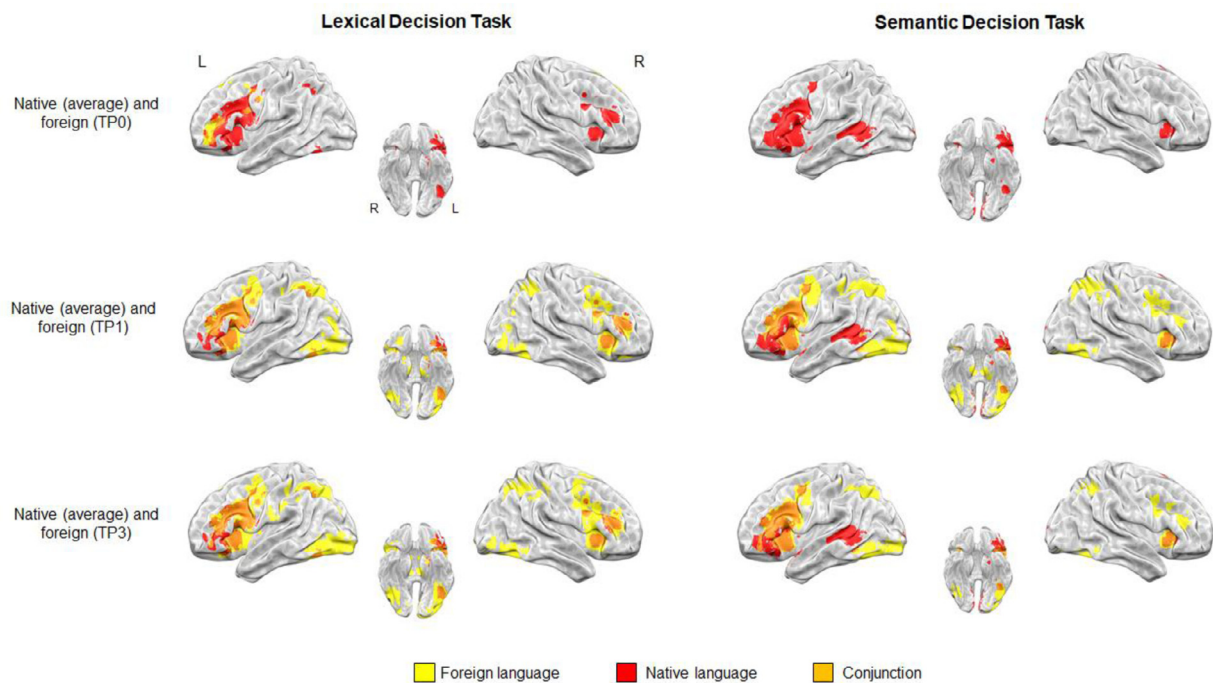


Fig. 5. Spatial convergence over time: whole-brain overlaps between native and foreign language networks. Activation maps for native language processing were averaged across all TPs. Left panel shows results from the lexical decision task at TP0 (upper), TP1 (middle) and TP3 (bottom). Right panel shows results from the semantic decision task. All maps represent the *experimental* > *control* contrasts from according tasks and are thresholded at $p < 0.05$ (FWEc). TP = time point; L = left hemisphere; R = right hemisphere.

right, middle and lower panels). The detailed lists of local maximas from the overlapping regions are presented in supplementary materials S1.4 (Tables S5 and S6).

Finally, we conducted a fine-tuned spatial convergence analysis of Dice coefficients calculated for anatomically-instructed regions. RMANOVAs revealed significant changes in foreign and native language overlaps in Broca's area in the lexical task (main effect of time $F(2.9, 49.7) = 3.2$; $p = 0.031$). Interestingly, post-hoc analyses indicated a temporal pattern different from the amplitude analysis (Fig. 3C), with gradual convergence increase until the end of training (TP3-TP0 $p = 0.002$, Bonferroni corrected) and a tendency for a return to baseline in the follow-up scan (Fig. 6). There were no significant effects of time in other ROIs in the lexical task (Visual Word Form Area $F(3.3, 55.6) = 1.3$; $p = 0.294$; left inferior parietal lobule $F(3.1, 51.9) = 2.7$; $p = 0.052$; anterior cingulate cortex $F(1.8, 30.9) = 1.9$; $p = 0.172$). No significant

changes over time were observed in the semantic task (Broca's area $F(3.3, 57.4) = 2$; $p = 0.117$; Visual Word Form Area $F(3.4, 57.2) = 0.8$; $p = 0.529$; left inferior parietal lobule $F(3.6, 60.5) = 1.9$; $p = 0.123$; anterior cingulate cortex $F(2.9, 50.2) = 0.8$; $p = 0.521$).

4. Discussion

Using a longitudinal design, we systematically observed functional reorganization in the reading network of beginning learners of Greek over a period of one year. We conducted two sets of analysis, the first with a focus on the amplitude of the fMRI signal, the second with respect to spatial convergence of this activity over time. In the first set, we tested the functional activity in response to lexical and semantic decision tasks in native and foreign language, at 5 different time points.

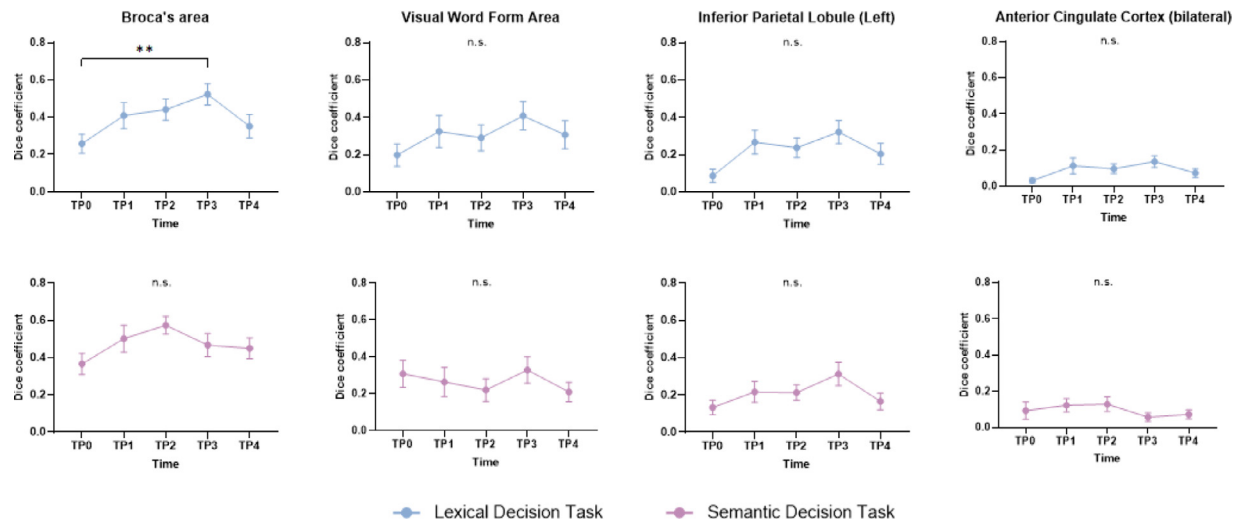


Fig. 6. Spatial convergence over time: Dice coefficients analysis in selected Regions of Interest. Data represent similarity between foreign and native language activation patterns calculated as the common number of voxels controlled by overall degree of activation in each condition (see Methods for details). ** $p < 0.01$, Bonferroni corrected; n.s. = not significant.

There was a common time course of activity in foreign language for both tasks with significant changes in brain activity occurring already in the first 3 months of instruction (between TP0 when subjects were naïve to the foreign language and TP1, the first scan after formal classroom instruction). These changes were found in the classic left-lateralized language network (defined after Dehaene, 2010) and included additional cognitive control areas, like the inferior parietal cortex (described as relevant in second language learning by Barbeau et al., 2017 and included in the network described by Abutalebi and Green, 2007). Parts of the inferior parietal cortex (inferior parietal sulcus) have also been linked to other processes relevant in reading pseudowords and letter-to-phoneme conversion (Dehaene, 2010; Cohen et al 2008; Ossmy et al 2014). Main effect of time for foreign language reading was driven by the initial 3 months of instruction, while no significant changes over time were revealed for native reading. Overall, our results are consistent with previous studies (Barbeau et al., 2017; Osterhout et al., 2008; Raboyeau et al., 2010).

Unlike Barbeau et al. (2017), we find differences between pre-exposure and post-exposure scans in an extended network, which includes more regions than the parietal lobe. We suggest that because our language of choice (Greek) is not written in Latin script, we avoid activation of the language network at TP0 as a response to written stimuli that might bear similarity to native words. Grapheme-to-phoneme conversion is a critical step in the development of reading skills (Blau et al., 2010) and might also be crucial to second language literacy (Brice et al., 2019).

Despite the differences between native and foreign alphabet, reading in the foreign language is associated with activity in regions that are part of the native reading network, such as the left inferior frontal gyrus and the left lateral occipital cortex (Liebig et al., 2017). We conclude that instead of building new representations of a foreign language, the native language network is recruited from early phases of foreign language reading onwards (Van de Putte et al., 2018).

Non-proficient reading of a foreign language requires more effort than reading of the mother tongue (Abutalebi et al., 2013; Del Maschio and Abutalebi, 2017) but that additional activity was shown to decrease when proficiency increases (Raboyeau et al., 2010; Stein et al., 2009). We observed this elevated activity for Broca's area, the visual word form area and the inferior parietal lobule (Fig. 4). We did not find a decrease in activity with time. Instead, the activity of the recruited areas was stable throughout the learning period and beyond. It could be argued that participants did not increase their proficiency level and hence the neu-

ral correlates of foreign reading remained stable. This is unlikely, since participants continued attending the Greek language course, where they were presented with increasingly difficult material. The behavioral task in the MR scanner was designed to be increasingly difficult at each session; performance is therefore not a direct indicator of proficiency levels in this study. Learning a new language is a demanding skill, and using a new language requires retrieving the acquired vocabulary from a dynamically expanding lexicon. Therefore, we created an ecologically valid task design by keeping task difficulty comparable across TPs. This approach differs from other longitudinal studies with a fixed set of vocabulary and might also explain some of the differences between this study and others.

Alternatively, the lacking decrease in activation of the foreign reading network could be explained by the constant task demand. It has previously been shown that non-cognate stimuli that do not resemble native words recruit a more extended network than cognate stimuli, even when they are well memorized (Raboyeau et al., 2010). It is therefore possible that the cognitive effort of retrieving information about given stimuli - rather than the overall proficiency level - are determining which networks are activated and to what extent. The reduced differences between native and foreign language processing point in the direction of neuronal reorganization occurring in response to foreign language learning. However, due to a lack of independent measures of proficiency and performance in the task, here, we cannot correlate these two measures with neural activity.

In the foreign lexical and semantic decision tasks, activity decreased between TP0 and TP1 in the bilateral angular gyrus. Since the angular gyrus is frequently discussed as a hub in the default mode network (Seghier, 2013), one explanation for the higher activity at TP0 in the foreign language is the character of the task at this time point. The instruction required participants to make a semantic or lexical decision, while they were not able to read the script. Participants have likely used a guessing strategy (reflected in the chance-level performance) which required little cognitive effort and allowed mind-wandering.

In both tasks, the spatial convergence of activity was extended after 3 months of instruction. Specifically, a larger part of inferior frontal gyrus was commonly activated by both languages at TP1 compared to TP0 in the lexical task. In the semantic task, the commonly activated areas extended from the supplementary motor cortex at TP0 to the bilateral early visual cortices, the left inferior frontal gyrus and the precentral gyrus. This confirms the observations made on the functional level. The additional overlap can be attributed to the additional activity observed

from TP1 onwards, and does not yet provide direct evidence for a shift in the network location.

For a more specific analysis of the similarity of the spatial activity maps, we computed dice coefficients. In the lexical task, the similarity of the activity pattern in Broca's area increased constantly between TP0 and TP3, resulting in a significantly higher similarity of native and foreign language at TP3 compared to TP0. This result suggests that differently from our observation of stable activity from TP1 onwards, the representation of the foreign language still changes over time. Assimilation of native and foreign language networks, especially in terms of spatial convergence could play a role in shaping the foreign language network, even before these changes become apparent in the functional data. Neuronal reorganization occurred rapidly, within 3 months (TP1), and resulted in elevated activity in the foreign language network when compared to the native network. We observe spatial convergence only in one specific task (lexical task) in the Broca's area. This convergence pattern does not follow the activity pattern of stable recruitment after TP1, but instead shows an increase in similarity till the end of the learning period. This makes it likely that a longer learning period, higher proficiency levels or a more immersive contact with the foreign language would strengthen this convergence, and that those changes would also become visible in the functional activity (Abutalebi et al., 2013; Grant et al., 2015; Stein et al., 2009). Due to a lack of independent measures of proficiency and performance in the tasks, we cannot correlate participants' individual proficiency with neural activity.

Cross-sectional studies comparing native and foreign reading have suggested that the orthographic transparency of the foreign language mediates the recruitment of dorsal or ventral regions (Oliver et al., 2016). The dorsal pathway involving the left opercular inferior frontal gyrus, the left superior temporal gyrus, and the left inferior parietal lobe is implicated in transparent languages that are read using a phonological decoding strategy. In contrast, the ventral pathway, consisting of the left triangular inferior frontal gyrus, the left orbital inferior frontal gyrus, and the left ventral occipital temporal lobe is recruited for opaque languages that rely on orthographic word-to-meaning mappings. In our study, we find activation in regions of both the dorsal pathway (left opercular inferior frontal gyrus, left inferior parietal lobe) and the ventral pathway (left triangular inferior frontal gyrus, left lateral occipital cortex). Neither Polish nor Greek are fully opaque, nor fully transparent languages (Protopapas and Vlahou, 2009; Schüppert et al., 2017), therefore, ventral and dorsal networks might be both involved.

The current study was conducted on women with higher education who spoke at least one second language. This study might therefore not generalize to first contact with a second language, as interactions between more than one non-native language have been reported (Pajak et al., 2016). We cannot exclude that the findings might also potentially vary based on gender, as women were shown to outperform men when learning phonologically familiar novel words (Kaushanskaya et al., 2011). However, a recent review shows much more similarities than differences between men and women in the brain activation during language processing (Sato, 2020).

Tasks used here were designed to elicit a generic response in the language network without making fine distinctions between sub-processes, such as pure phonological processing in pseudo-words and phonologic-semantic processing in meaningful words. Following studies could investigate phonological, semantic, lexical and orthographic networks separately, as has been done in children (Liebig et al., 2017) to identify dynamics of different processing streams during foreign language acquisition. One prediction based on the presented data and Brice et al (2019) would be that the first neuroplasticity changes should occur in structures associated with grapheme-phoneme conversion (ventral occipito-temporal cortex, middle temporal gyrus, posterior parietal cortex, inferior frontal gyrus and supplementary motor cortex - according to Liebig et al., 2017).

The task instruction was consistent throughout the course of the study. In the first session of the foreign language, participants therefore guessed the correct answer. Behaviorally, we observed poorer per-

formance at TP0, along with faster reaction times. Although activity differences between TP0 and TP1 could result from the decrease in reaction times, it is more likely that both changes in RTs and activations are related to a strategy shift.

Attendance of participants decreased during the course of the experiment (see Fig. 1A, blue dots). The average time of classroom instruction was therefore not equal between scanning sessions. Due to homework that was solved outside of the classroom, and self-paced learning, we cannot estimate the full time of exposure to Greek between scanning sessions.

We included a break between TP3 and TP4 without courses, in which the activity pattern remains largely unaffected. The irregular attendance might result in slower learning rates and lower overall proficiency of the participants than expected from regular attendance. However, given that the observed network is still stable after a break of several months, this does not challenge the observation that stable recruitment of task-relevant areas happens rapidly within the first three months. Last but not least, given the rapid change observed within the first months of learning, one could wonder if this does indeed reflect a sudden switch-on of the language network or if this change can be observed in gradual recruitment of sub-regions in a finer temporal resolution.

In summary, our data shows that most crucial changes in the foreign reading network happen with the transition from mere physical stimulation (unknown script) to lexical and semantic representations of the word or phrase. These changes include switching the task solving strategy from guessing to making informed decisions that is reflected in task performance and happen rapidly within the first months of instruction. Classical reading and cognitive control areas are recruited to a stronger extent than when the same task is solved in the mother tongue. A rapid recruitment of the language network for foreign language processing has been reported before (McLaughlin et al., 2004; Raboyeau et al., 2010). We extend these findings by showing that after an initial "switch-on" of the language system, foreign reading recruits a stable network, given that the task demands are kept constant. This offers the interesting perspective that the recruitment of additional regions for foreign language processing is mediated by task demand and cognitive effort rather than by proficiency. In the present data set, no evidence is found for an assimilation of native and foreign reading networks over 6 months of classroom learning.

Declaration of Competing Interest

The authors declare no competing financial interests

Acknowledgements

This study was supported by a National Science Centre Poland grant (2014/14/M/HS6/00918) awarded to A.M.

C.K. received the scholarship „Metropolen in Osteuropa“ from Studienstiftung des deutschen Volkes and Alfred Krupp von Bohlen und Halbach-Stiftung.

J.M. was additionally supported National Science Centre Poland grant (2017/27/N/HS6/02669) and ETIUDA grant (2019/32/T/HS6/00496)

A.B. was additionally supported National Science Centre Poland grant (2017/27/N/HS6/02722) and ETIUDA grant (2019/32/T/HS6/00529)

The study was conducted with the aid of CePT research infrastructure purchased with funds from the European Regional Development Fund as part of the Innovative Economy Operational Programme, 2007–2013.

We would like to thank Natalia Sokólska for co-organizing Greek course. Dawid Drożdziel and Bartosz Kossowski for technical support.

Credit author statement

Clara Kuper: Software, Formal analysis, Data curation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Jacek Matuszewski: Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing - Review & Editing, Visualization.

Anna Banaszkiewicz: Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing.

Łukasz Bola: Methodology, Writing - Review & Editing.

Michał Szczepaniak: Software, Investigation, Data curation.

Małgorzata Draps: Resources, Investigation.

Przemysław Kordos: Methodology, Resources, Investigation.

Marcin Szwed: Conceptualization, Methodology, Writing - Review & Editing.

Katarzyna Jednoróg: Conceptualization, Methodology, Formal analysis, Resources, Writing - Original Draft, Writing - Review & Editing.

Artur Marchewka: Conceptualization, Methodology, Formal analysis, Resources, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Data/code availability statement

Thresholded maps for each 2nd level contrast and behavioral performance are available for downloads: https://lobi.nencki.gov.pl/a/2-greek_supplementary.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.neuroimage.2020.117544](https://doi.org/10.1016/j.neuroimage.2020.117544).

References

- Abutalebi, J., Annoni, J.M., Zimine, I., Pegna, A.J., Seghier, M.L., Lee-Jahnke, H., Khateb, A., 2008. Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb. Cortex* 18 (7), 1496–1505. doi:[10.1093/cercor/bhm182](https://doi.org/10.1093/cercor/bhm182).
- Abutalebi, J., Della Rosa, P.A., Ding, G., Weekes, B., Costa, A., Green, D.W., 2013. Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex* 49 (3), 905–911. doi:[10.1016/j.cortex.2012.08.018](https://doi.org/10.1016/j.cortex.2012.08.018).
- Abutalebi, J., Green, D., 2007. Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20. doi:[10.1016/j.jneuroling.2006.10.003](https://doi.org/10.1016/j.jneuroling.2006.10.003).
- Ashburner, J., Ridgway, G.R., 2013. Symmetric diffeomorphic modeling of longitudinal structural MRI. *Front. Neurosci.* 6 (FEB), 1–19. doi:[10.3389/fnins.2012.00197](https://doi.org/10.3389/fnins.2012.00197).
- August, D., Shanahan, T. (Eds.), 2006. *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Lawrence Erlbaum Associates; The Center for Applied Linguistics.
- Barbeau, E.B., Chai, X.J., Chen, J.-K., Soles, J., Berken, J., Baum, S., Klein, D., 2017. The role of the left inferior parietal lobule in second language learning: an intensive language training fMRI study. *Neuropsychologia* 98 (August 2016), 169–176. doi:[10.1016/j.neuropsychologia.2016.10.003](https://doi.org/10.1016/j.neuropsychologia.2016.10.003).
- Bennet, C.M., Miller, M.B., 2010. How reliable are the results from functional magnetic resonance imaging? *Ann. N. Y. Acad. Sci.* 1191, 133–155. doi:[10.1111/j.1749-6632.2010.05446.x](https://doi.org/10.1111/j.1749-6632.2010.05446.x).
- Blau, Vera, Reithler, Joel, van Attefeldt, Nienke, Seitz, Jochen, Gerretsen, Patty, Goebel, Rainer, Blomert, Leo, 2010. Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain* 133 (3), 868–879. doi:[10.1093/brain/awp308](https://doi.org/10.1093/brain/awp308).
- Brett, M., Anton, J.-L., Valabregue, R., Jean-Baptiste, P., 2000. Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan. Available on CD-ROM in NeuroImage, Vol 16, No 2.
- Brice, H., Mencl, W.E., Frost, S.J., Bick, A.S., Rueckl, J.G., Pugh, K.R., Frost, R., 2019. Neurobiological signatures of L2 proficiency: evidence from a bi-directional cross-linguistic study. *J. Neurolinguistics* 50 (February), 7–16. doi:[10.1016/j.jneuroling.2018.02.004](https://doi.org/10.1016/j.jneuroling.2018.02.004).
- Chyl, K., Kossowski, B., Debska, A., Łuniewska, M., Banaszkiewicz, A., Zelechowska, A., Jednoróg, K., 2018. Prereader to beginning reader: Changes induced by reading acquisition in print and speech brain networks. *J. Child Psychol. Psychiatry* doi:[10.1111/jcpp.12774](https://doi.org/10.1111/jcpp.12774).
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., Montavont, A., 2008. Reading normal and degraded words: contribution of the dorsal and ventral visual pathways. *Neuroimage* 40, 353–366.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: a Dual Route Cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi:[10.1037/0033-295X.108.1.204](https://doi.org/10.1037/0033-295X.108.1.204).
- Das, T., Padakannaya, P., Pugh, K.R., Singh, N.C., 2011. Neuroimaging reveals dual routes to reading in simultaneous proficient readers of two orthographies. *Neuroimage* 54 (2), 1476–1487. doi:[10.1016/j.jacc.2007.01.076](https://doi.org/10.1016/j.jacc.2007.01.076).
- Dehaene, S., 2010. *Reading in the Brain*. Penguin, Viking.
- Del Maschio, N., Abutalebi, J., 2017. Neurobiology of bilingualism. In: *Bilingual Cognition and Language, The state of the Science Across Its Subfields*, pp. 332–352.
- DeLuca, V., Rothman, J., Bialystok, E., Platsikas, C., 2019. Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proc. Natl. Acad. Sci.* 116 (15), 7565–7574. doi:[10.1073/pnas.1811513116](https://doi.org/10.1073/pnas.1811513116).
- Draganski, B., Christian, G., Volker, B., Gerhard, S., Ulrich, B., Arne, M., 2004. Changes in grey matter induced by training. *Nature* 427 (6972). doi:[10.1038/427312a](https://doi.org/10.1038/427312a), 312–312.
- Draganski, B., Kherif, F., Lutti, A., 2014. Computational anatomy for studying use-dependent brain plasticity. *Front. Human Neurosci.* 8 (June), 1–7. doi:[10.3389/fn-hum.2014.00380](https://doi.org/10.3389/fn-hum.2014.00380).
- Ghazi-Saidi, L., Ansaldi, A.I., 2017. The neural correlates of semantic and phonological transfer effects: language distance matters. *Bilingualism* 20 (5), 1080–1094. doi:[10.1017/S136672891600064X](https://doi.org/10.1017/S136672891600064X).
- Grainger, J., Midgley, K., & Holcomb, P. J. (2008). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). <https://doi.org/10.1075/btl.128.02gar>
- Grant, A.M., Fang, S.Y., Li, P., 2015. Second language lexical development and cognitive control: a longitudinal fMRI study. *Brain Lang.* 144, 35–47. doi:[10.1016/j.bandl.2015.03.010](https://doi.org/10.1016/j.bandl.2015.03.010).
- Green, D.W., 2003. Neural basis of lexicon and grammar in L2 acquisition: the convergence hypothesis. The Interface between Syntax and the Lexicon in Second Language Acquisition 197–208. doi:[10.1075/iald.30.10gre](https://doi.org/10.1075/iald.30.10gre).
- Green, D.W., Price, C.J., Crinion, J., 2006. Convergence, degeneracy and control. *Language Learning* 56, 1–16.
- Holmes, D., Moody, P., Dine, D., Trueman, L., 2016. Research Methods for the Biosciences, 3rd ed. Retrieved from <https://global.oup.com/academic/product/research-methods-for-the-biosciences-9780198728498?cc=pl&lang=en&#>.
- Jacobs, A.M., Ziegler, J.C., 2015. Neurocognitive psychology of visual word recognition. In: Wright, J.D. (Ed.), *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, Oxford, pp. 214–219.
- Kaushanskaya, Margarita, Marian, Viorica, Yoo, Jeewon, 2011. Gender differences in adult word learning. *Acta Psychologica* 137 (1), 24–35. doi:[10.1016/j.actpsy.2011.02.002](https://doi.org/10.1016/j.actpsy.2011.02.002).
- Kelly, A.M.C., Garavan, H., 2005. Human functional neuroimaging of brain changes associated with practice. *Cereb. Cortex* 15 (8), 1089–1102. doi:[10.1093/cercor/bhi005](https://doi.org/10.1093/cercor/bhi005).
- Kim, S.Y., Qi, T., Feng, X., Ding, G., Liu, L., Cao, F., 2015. How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean-Chinese-English trilinguals. *Neuroimage* 129, 25–39. doi:[10.1016/j.neuroimage.2015.11.068](https://doi.org/10.1016/j.neuroimage.2015.11.068).
- Lerma-Usabiaga, G., Carreiras, M., Paz-Alonso, P.M., 2018. Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *PNAS* 115 (42), E9981–E9990. doi:[10.1073/pnas.1803003115](https://doi.org/10.1073/pnas.1803003115).
- Li, P., Legault, J., Litcofsky, K.A., 2014. Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58 (June), 301–324. doi:[10.1016/j.cortex.2014.05.001](https://doi.org/10.1016/j.cortex.2014.05.001).
- Liebig, J., Froehlich, E., Morawetz, C., Braun, M., Jacobs, A.M., Heekeren, H.R., Ziegler, J.C., 2017. Neurofunctionally dissecting the reading system in children. *Develop. Cogn. Neurosci.* 27 (August), 45–57. doi:[10.1016/j.dcn.2017.07.002](https://doi.org/10.1016/j.dcn.2017.07.002).
- Liu, H., Cao, F., 2016. L1 and L2 processing in the bilingual brain: a meta-analysis of neuroimaging studies. *Brain Lang.* 159, 60–73. doi:[10.1016/j.bandl.2016.05.013](https://doi.org/10.1016/j.bandl.2016.05.013).
- Mandera, P., Keuleers, E., Wodniecka, Z., Brysbaert, M., 2015. Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behav. Res. Methods* 47 (2), 471–483. doi:[10.3758/s13428-014-0489-4](https://doi.org/10.3758/s13428-014-0489-4).
- McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. Neurosci.* 7 (7), 703–704. doi:[10.1038/nrn1264](https://doi.org/10.1038/nrn1264).
- Meschyan, G., Hernandez, A.E., 2006. Impact of language proficiency and orthographic transparency on bilingual word reading: an fMRI investigation. *Neuroimage* 29 (4), 1135–1140. doi:[10.1016/j.neuroimage.2005.08.055](https://doi.org/10.1016/j.neuroimage.2005.08.055).
- Morgan-Short, K., Finger, I., Grey, S., Ullman, M.T., 2012. Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One* 7 (3). doi:[10.1371/journal.pone.0032974](https://doi.org/10.1371/journal.pone.0032974).
- Oliver, M., Carreiras, M., Paz-Alonso, P.M., 2016. Functional dynamics of dorsal and ventral reading networks in bilinguals. *Cereb. Cortex* (December) 5431–5443. doi:[10.1093/cercor/bhw310](https://doi.org/10.1093/cercor/bhw310).
- Ossmy, O., Ben-Shachar, M., Mukamel, R., 2014. Decoding letter position in word reading. *Cortex* 59, 74–83.
- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Hirschensohn, J., 2008. Second-language learning and changes in the brain. *J. Neurolinguistics* 21 (6), 509–521. doi:[10.1016/j.jneuroling.2008.01.001](https://doi.org/10.1016/j.jneuroling.2008.01.001).
- Pajak, B., Fine, A.B., Kleinschmidt, D.F., Jaeger, T.F., 2016. Learning additional languages as hierarchical probabilistic inference: insights from first language processing. *Language Learning* 66 (4), 900–944. doi:[10.1111/lang.12168](https://doi.org/10.1111/lang.12168).
- Perfetti, C.A., Liu, Y., 2005. Orthography to phonology and meaning: comparisons across and within writing systems. *Reading and Writing* 18 (3), 193–210. doi:[10.1007/s11145-004-2344-y](https://doi.org/10.1007/s11145-004-2344-y).
- Protopapas, A., Vlahou, E.L., 2009. A comparative quantitative analysis of Greek orthographic transparency. *Behav. Res. Methods* 41 (4), 991–1008. doi:[10.3758/BRM.41.4.991](https://doi.org/10.3758/BRM.41.4.991).
- Raboyeau, G., Marcotte, K., Adrover-Roig, D., Ansaldi, A.I., 2010. Brain activation and lexical learning: the impact of learning phase and word type. *Neuroimage* 49 (3), 2850–2861. doi:[10.1016/j.neuroimage.2009.10.007](https://doi.org/10.1016/j.neuroimage.2009.10.007).
- Rueckl, J.G., Paz-Alonso, P.M., Molfese, P.J., Kuo, W.-J., Bick, A., Frost, S.J., ..., Frost, R., 2015. Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proc. Natl. Acad. Sci.* 112 (50), 15510–15515. doi:[10.1073/pnas.1509321112](https://doi.org/10.1073/pnas.1509321112).
- Sato, M., 2020. The neurobiology of sex differences during language processing in healthy adults: A systematic review and a meta-analysis. *Neuropsychologia* 140, 107404.

- Schüppert, A., Heeringa, W., Golubović, J., Gooskens, C., 2017. Write as you speak? A cross-linguistic investigation of orthographic transparency in 16 Germanic, Romance and Slavic languages. *From Dialectometry to Semantics* 303–313.
- Siuda-Krzywicka, K., Bola, Ł., Paplińska, M., Sumera, E., Jednoróg, K., Marchewka, A., Szwed, M., 2016. Massive cortical reorganization in sighted braille readers. *ELife* 5 (MARCH2016), 1–26. doi:[10.7554/eLife.10762](https://doi.org/10.7554/eLife.10762).
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehmann, C., Wiest, R., Dierks, T. (2009). Reduced frontal activation with increasing 2nd language proficiency. 47, 2712–2720. <https://doi.org/10.1016/j.neuropsychologia.2009.05.023>.
- Szwed, M., Qiao, E., Jobert, A., Dehaene, S., Cohen, L., 2014. Effects of literacy in early visual and occipitotemporal areas of Chinese and French readers. *J. Cogn. Neurosci.* 26 (3), 459–475.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15 (1), 273–289. doi:[10.1006/nimg.2001.0978](https://doi.org/10.1006/nimg.2001.0978).
- Van de Putte, E., De Baene, W., Price, C.J., Duyck, W., 2018. Neural overlap of L1 and L2 semantic representations across visual and auditory modalities: a decoding approach. *Neuropsychologia* 113 (March), 68–77. doi:[10.1016/j.neuropsychologia.2018.03.037](https://doi.org/10.1016/j.neuropsychologia.2018.03.037).
- Zatorre, R.J., Fields, R.D., Johansen-Berg, H., 2012. Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nat. Neurosci.* 15 (4), 528–536. doi:[10.1038/nn.3045](https://doi.org/10.1038/nn.3045).