Published in final edited form as:

Neuropsychologia. 2014 September ; 62: 77–86. doi:10.1016/j.neuropsychologia.2014.07.007.

Dynamic Changes in Network Activations Characterize Early Learning of a Natural Language

Elena Plante¹, Dianne Patterson¹, Natalie S. Dailey¹, Kyle, R. Almyrde¹, and Julius Fridriksson²

¹Department of Speech, Language, and Hearing Sciences, The University of Arizona, AZ

²Department of Communication Sciences and Disorders, The University of South Carolina, SC

Abstract

Those who are initially exposed to an unfamiliar language have difficulty separating running speech into individual words, but over time will recognize both words and the grammatical structure of the language. Behavioral studies have used artificial languages to demonstrate that humans are sensitive to distributional information in language input, and can use this information to discover the structure of that language. This is done without direct instruction and learning occurs over the course of minutes rather than days or months. Moreover, learners may attend to different aspects of the language input as their own learning progresses. Here, we examine processing associated with the early stages of exposure to a natural language, using fMRI. Listeners were exposed to an unfamiliar language (Icelandic) while undergoing four consecutive fMRI scans. The Icelandic stimuli were constrained in ways known to produce rapid learning of aspects of language structure. After approximately 4 minutes of exposure to the Icelandic stimuli, participants began to differentiate between correct and incorrect sentences at above chance levels, with significant improvement between the first and last scan. An independent component analysis of the imaging data revealed four task-related components, two of which were associated with behavioral performance early in the experiment, and two with performance later in the experiment. This outcome suggests dynamic changes occur in the recruitment of neural resources even within the initial period of exposure to an unfamiliar natural language.

Keywords

language; statistical learning; unguided learning; functional connectivity; fMRI

© 2014 Elsevier Ltd. All rights reserved.

Corresponding author: E. Plante, PhD Department of Speech, Language, & Hearing Sciences, PO Box 210071, The University of Arizona, Tucson, AZ 85721-0071, eplante@email.arizona.edu, Phone: +001 520-621-5080.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Although the neural regions recruited to support language processing are now well described (Cappa, 2012; Price, 2010; 2012), far less is known about the emergence of the cortical language network. It is likely that the process of learning language recruits at least some fundamentally different neural resources than the ones used once language is acquired. For example, on first exposure to a language, learners often have difficulty distinguishing individual words within conversational speech. When a language has been learned, listeners have little difficulty hearing individual words and can recognize semantic, morphologic, and syntactic features of the language as well. Therefore, it is likely that neural resources that support skills such as the segmentation of words from running speech might be actively engaged early on, but not after the language has been acquired.

Importantly, languages acquired in naturalistic contexts are typically learned from exposure to the language, and explicit feedback is rare. This type of unguided language learning can be differentiated from explicit teaching in that learners are not given information concerning the nature of the language or how it is structured. Unguided learning can also be differentiated from classification learning in that learners do not receive regular feedback that shapes their conceptualization of the input. This latter distinction is important because the presence of feedback during learning appears to alter the brain regions recruited during learning (see Poldrack & Rodriguez, 2004 for a review). In particular, the presence and type of feedback during learning of tasks involving artificial grammars has been associated with changes in the electrophysiologic response (Opitz, Ferdinand, & Meckinger, 2011) and appears to engage systems associated with application of memory strategies (Fletcher, Büchel, Josephs, Friston, & Dolan, 1999).

In this study, we examine unguided language learning of a natural language (Icelandic) in order to determine how the language network changes during the earliest period of exposure to an unfamiliar language. Although others have focused on mapping meaning from context during exposure to an unfamiliar language (Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010), here we focus specifically on learning of the structure of the language, independent of its meaning. However, like the Veroude et al study, we also focused on adult learners. A study of adults, as opposed to children, allows for a view of activation changes associated with learning that is independent of the changes associated with maturation (e.g., Plante, Holland, & Schmithorst, 2006; Schmithorst, Holland, & Plante, 2007; Szaflarski, Altaye, Rajagopal, Eaton, Meng, Plante, & Holland, 2012; Szaflarski, Schmithorst, Altaye, Byars, Rett, Plante, & Holland, 2006). In children who are acquiring language, these two constructs can be difficult to dissociate.

Most imaging studies of unguided language learning have employed highly constrained artificial languages rather than natural language stimuli. Most have focused on the process of segmenting pseudowords from continuous speech, based on the paradigms of Saffran, Aslin, Newport and colleagues (Aslin, Saffran, & Newport, 1998; Saffran, Newport, & Aslin, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) that first described this process behaviorally. These behavioral studies established that learners (infants and adults) extract words from continuous speech by tracking the statistical distribution of phoneme-

level information. The ability to segment speech into individual words is a critical first step in language learning, in that individual words must be identified before meaning can be mapped or the syntactic relations among words can be learned. The fMRI studies that have examined word segmentation describe activation in the superior and middle temporal gyrus. (Cunillera, Càmarà, Toro, Marco-Pallares, Sebastian-Galles, Ortiz, & Rodriguez-Fornells, 2009; Karuza, Newport, Aslin, Starling, Tivarus, & Bavelier, 2013; McNealy, Mazziotta, & Dapretto, 2006), as well as dorsolateral prefrontal cortex (Cunillera et al., 2009; McNealy et al., 2006) and postcentral gyrus (Karuza et al., 2013). However, there are differences among studies in terms of the direction of change over time across studies (Cunillera et al., 2009; McNealy et al., 2006). Posterior temporal activation, extending into the supramarginal gyrus decreased after 2 minutes of exposure in the Cunillera et al study, but increased over the course of an 8 minute scan used in the McNealey et al study. Decreases in activation with time were also reported for dorsolateral prefrontal activation (Cunillera et al., 2009) and ventromedial frontal activation (McNealey et al., 2006).

Learners also track predictable relations among the newly-identified words. This transition between identifying individual words and the relations among words has been captured in an event-related potential study of artificial language (De Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Levi, 2007). This study demonstrated that an electrophysiologic response (the P2 waveform) to structural relations among 3-word combinations in an artificial language occurred after the emergence of response associated with recognition of word forms (N400 waveform). Examination of unguided learning of a more complex finite state grammar was undertaken by Newman-Norlund Frey, Petitto, & Grafton, 2006. Their subjects were trained on a set of novel vocabulary arranged to reflect syntactic rules for combining words. These learners were scanned during their initial exposure to the grammar and again after extensive out-of-scanner training on the grammar. Unlike results for the word segmentation studies described above, Newman-Norlund et al reported increased activation in the left inferior frontal gyrus accompanied by decreased activation in this region in the right hemisphere.

Both of these studies employed micropauses between words. Micropauses, which are not perceptually obvious to the listener, are known to facilitate attention to dependent relations among words (cf. Peña, Bonatti, Nespor, & Mehler, 2002). These pauses also served to distinguish the boundaries between individual words, which would not necessarily be temporally separated in naturally occurring speech. However, such pauses could also function as a type of prosodic cue to language structure. When both statistical and prosodic cues are presented in concert, reduced activation is noted in frontal regions compared to presentation of statistical cues in the absence of stress cues (McNealy Mazziotta, & Dapretto, 2010) suggesting that the presence of multiple converging cues to language structure reduces the processing load. Therefore, it appears that when listeners are allowed to use multiple converging cues to language structure, processing is different than when the input is constrained in such a way that use of only a single cue is possible. In addition, change over time may be particularly pronounced when a learning context allows for the extraction of multiple aspects of language structure (i.e., word forms, prosodic structure, grammar) in that listeners may weigh the value of these cues differently as learning progresses (cf., Theissen & Saffran 2003). Conversely, other evidence suggests that adult

learners are able to track multiple word-level distributional properties of the input at the same time (Romberg & Saffran, 2013). These types of studies suggests that examination of unguided natural language learning, with the presence of multiple cues to language structure, would provide an important complement to the more constrained paradigms offered by artificial language studies.

Previous fMRI studies of learning have relied on general linear model (GLM) analyses. However, it is possible to extract additional information about the fMRI signal than is permitted by GLM models. By definition, GLM analysis uses a pre-specified “general” model and is unable to detect true differences in the temporal aspects of the hemodynamic response that may occur across spatially distributed brain regions. Independent Component Analysis (ICA) offers an alternate analysis approach that is sensitive to this source of signal variation.

ICA is a model-free method of detecting distinct (i.e., independent) sources of signal variation within the complex BOLD response. ICA offers several advantages for analysis of fMRI data in general and for the study of learning in particular. First, ICA can be more sensitive to true signal variation than GLM because ICA decomposes the complex BOLD signal into a set of independent signals that are otherwise combined by a GLM analysis. In other words, sources of signal variation detected by ICA cannot be differentiated from error variance in GLM. Accordingly, because ICA accounts for multiple sources of true variance, the remaining variance (error variance) is reduced relative to the GLM. Furthermore, signal related to physiologic noise or subject movement can be separated from task-related signal because these two signal sources are often statistically (and physiologically) independent. Finally, the model-free nature of ICA allows for the prospect that signal associated with different aspects of processing may diverge from an omnibus model of the hemodynamic response. Because these features of ICA tend to increase statistical power, ICA is often a more sensitive technique than the computationally-simpler GLM approach, and consequently typically reveals more regions of statistically significant activation (e.g., Karunanayaka, Schmithorst, Vannest, Szaflarski, Plante, & Holland, 2010; Schmithorst, Holland, & Plante, 2006; 2007).

This segregation of the BOLD signal into component signals has two other important effects that may be particularly relevant for the study of dynamic processes involved in unguided learning. First, differences related to the temporal aspect of the signal can refine interpretation of a distributed neural network. Regions that activate on the same time-course (i.e., a single IC) are presumed to be functionally connected. This implies that co-activating regions operate as a subnet of regions that can be distinguished from other subnets (i.e., other ICs) within a larger task-related neural network. Second, changes in the strength of activation over time of an IC and its association with behavioral performance can constrain explanations of how the regions of interest contribute to the learning process.

In this study, we use ICA analysis to examine processing of a natural language over the period of four consecutive fMRI scans. Participants were exposed to Icelandic sentences that were composed using a small vocabulary set. Participants were tested concerning their ability to distinguish legal Icelandic sentences from illegal strings that used the same

vocabulary set arranged with one or more word-order violations. Consistent with natural language learning, participants were not given feedback concerning their performance.

Our over-arching hypothesis is that dynamic change would characterize the regions that activate over time as participants progress from their first exposure to unfamiliar running speech. More specifically, we expect that regional activation previously associated different aspects of learning in artificial language paradigms will reflect different signal time courses and therefore separate independent components. Specifically, ICA will segregate activation in posterior temporal / supramarginal regions previously reported for word segmentation tasks from activation in inferior frontal regions reportedly critical for finite state grammar learning. Moreover, to the extent that these learning processes are differentially engaged over time, these different ICs will correlate with behavioral performance at different points in time within the experiment. We also expect that ICs in regions associated with more general processing capacities (e.g., dorsolateral prefrontal, inferior parietal) will be independent of ICs implicated in statistical learning of language.

2. Materials and methods

2.1. Subjects

Fourteen subjects, 8 males and 6 females (age range 19 - 28, median = 24) participated in the study. Age was not reported for two of the participants; however they were within the age range of the other participants. All participants spoke English as their primary language. Two spoke additional languages (Arabic, French, Italian). Prior to scanning, potential participants were screened and excluded from the study if there was a history of neurological disorders. We also excluded one potential participant who reported prior exposure to Icelandic. Informed consent was obtained for all participants prior to participating.

2.2. Materials

The experiment utilized a block design with three conditions (listen, test, and control). Participants listened to stimuli consisting of Icelandic sentences during the experiment. The stimuli are listed, in the order in which they were presented during the experiment, in the Appendix. Sentences were constructed around a limited set of vocabulary consisting of seven nouns, seven verbs, eight adjectives, two adverbs and three prepositions. This limited number of unique content words presented during the experiment was intended to facilitate learning based on how the items could be combined in syntactically legal Icelandic sentences. This limited lexicon also constrains the phoneme level-statistics that demark individual words within the speech stream. Icelandic and English phonology partially overlap, which should assist English speakers to quickly detect distributional information within the speech stream. However, word production by a native Icelandic speaker clearly sounds foreign to an unfamiliar listener.

Icelandic words primarily conform to a trochaic stress pattern, which is an additional cue to word boundaries within the speech stream. English also uses trochaic stress, but to a lesser extent than Icelandic. The presence of a trochaic stress pattern may assist learners in using

stress as a cue to word boundaries. In addition, spoken sentences contain phrase-level prosodic information that could assist a listener in distinguishing between syntactically legal and illegal sentences, in that illegal sentences are more difficult to produce with a completely typical stress pattern.

Unlike English, Icelandic uses case marking of nouns and adjectives, thus creating distinctive two and three syllable words that appeared in different sentence positions. Therefore, case marking provided a third cue, in addition to word order and sentential stress patterns, to syntactic structure.

All sentences were spoken in a natural voice and prosody by a native male Icelandic speaker. Sentences were digitally recorded and edited using Sound Forge 7.0 (Sony, 2003) to fit the parameters of the fMRI scan. Sets of six sentences were presented in a block, with an inter-sentence interval of approximately 700ms. Inter-sentence intervals were adjusted slightly so that each six-sentence block had a duration of exactly 21.6 seconds. Therefore, the total exposure to legal Icelandic sentences during the listening blocks of each scan was 2 minutes, 9.6 seconds.

The test block was composed of sentences that were either grammatically correct or grammatically incorrect. Both correct and incorrect sentences were read by the Icelandic speaker using an intonation pattern appropriate for the sentence frame. There were 18 correct and 18 incorrect sentences in total and these were randomly distributed in blocks of 6 sentences each. Grammatical sentences were identical to those presented during the listening blocks (but not necessarily the immediately preceding block). In contrast, ungrammatical sentences included the same set of vocabulary items, but were re-arranged to form an ungrammatical order, not only relative to the grammatical items heard within the context of the experiment, but also relative to the Icelandic language in general. Therefore, recognition of the specific ordering the words, rather a sense of having heard each of the words previously, was important in terms of being able to differentiate correct vs. incorrect test items. An inter-stimulus interval, ranging from 1.6 - 2.0 s., separated each sentence. This interval allowed time for the participant to indicate whether the sentence was grammatical or not. This response time was based on pilot testing. The total length of each test block was 21.6 seconds.

Control blocks were composed of six strings of words that were clipped from the listening blocks and each Icelandic syllable was temporally reversed. Syllables were then re-assembled into sentence-length strings. The inter-stimulus interval was approximately 700 ms to mimic that of the listening block. The total length of each control block was 21.6 seconds.

In addition to the listening and test stimuli, verbal cues to the participants were also recorded. These were meant to ensure that the participant understood what was expected prior to the onset of each block within the scan. The listen block cue ('Listen') lasted 2.4 ms, whereas the test block cue ('Is this right?') and control block cue ('Stop') each had a duration of 7.2ms. All cues were followed by a silent period so that the entire cue period length was 2.4 seconds (1 TR period). A female native English speaker recorded cues

digitally so that the voice for the cue was perceptually distinct from the male Icelandic speaker.

2.3 Procedures

Prior to scanning, participants were provided with a practice set that consisted of the same stimuli used during the four scans. Therefore, all participants had 2 minutes and 9.6 seconds of exposure to Icelandic prior to the first scan. Pilot data indicated that some, if not most, participants would need this prescan input to achieve learning within the time period of the experiment. In addition, it ensured that all participants understood the nature of the task prior to the first scan.

The experiment used the same stimuli across all four scans. Stimuli were presented through MR-compatible headphones (Resonance Technologies), which provided some attenuation of the background scanner noise. Sound levels were adjusted so that participants subjectively reported stimuli were audible, but not uncomfortably loud. Stimuli were arranged in a block design with six alternating blocks of listen, test and control stimuli. Cues to the participant were presented prior to each block, indicating the nature of the upcoming block. During test blocks participants indicated whether the sentences heard were correct or incorrect by pressing a button with either their right or left index finger. The hand assigned to 'yes' and 'no' response buttons was counterbalanced across subjects. The same stimuli were repeated over the course of four consecutive scans. After each scan, participants were asked to rate how difficult they perceived the task to be. A five point rating scale was used where 1 corresponded to very easy and 5 corresponded to very difficult.

2.4. Data acquisition

Scans were acquired using a 1.5-Tesla G.E. Signa magnet with a standard head coil. The functional scans consisted of 22 5-mm slices with a 1-mm inter-slice gap interval obtained in the axial plane. A single shot spiral acquisition protocol was used (Glover & Law, 1999). The repetition time [TR] was 2.4s with an echo time [TE] of 40ms. The field of view [FOV] was 22 cm² with a matrix size of 64 × 64 pixels. Spatial resolution for the functional scans was 3.44 mm × 3.44 mm × 5 mm. There were 214 time points for the experimental portion of the scan, with 4 additional time points at the beginning of each scan to allow the magnet to achieve equilibrium. Each functional scan lasted a total of 8 minutes, 43.2 seconds.

Two structural scans were obtained. A high-resolution SPGR gradient-echo pulse sequence (spoiled GRASS) was used for visualizing activation for individual subjects. The field of view [FOV] was 151 cm² with a matrix of 256 × 256, NEX = 2, 124 contiguous 1.5 mm slices. A fast spin echo [FSE] was obtained with a TR of 500 and a TE of 14 (Flip = 90/180, FOV = 22 cm², Matrix = 256 X 256, Number of Excitations [NEX] = 2). This scan had the same slice thickness and alignment as with the functional scans. Therefore, this structural scan, which had a higher resolution than the functional scans, could be used to ensure that the best alignment possible was achieved between the low resolution functional scan and the high resolution SPGR.

2.5 Image analysis

Functional images were analyzed using both GLM and ICA procedures. As the GLM analysis is not central to this study, we present it in the supplementary materials. Preprocessing of functional image data was accomplished using AFNI software (version AFNI_2011_12_21_1014, Cox, 2012). Raw images were evaluated to determine that the image intensity had reached equilibrium prior to the start of the behavioral stimuli and pre-stimulus scans were removed from each image set. Images were registered to a base image selected for each subject to represent minimal deviation in signal intensity from the average obtained for each time point across the scan period. The degree of head movement was evaluated for each participant, and these data were retained for use as a covariate in the GLM analysis. Slice data was time-shifted, using Fourier interpolation, so that all slices were equated to the time of a middle slice. Each scan was despiked, using AFNI's algorithm and then normalized to a scanner-specific template and fit into standard space. The fMRI data were smoothed using a Gaussian kernel (7 mm, FWHM).

The GLM analysis was performed on the preprocessed data using AFNI software. AFNI's standard hemodynamic response model was convolved with the box-car function corresponding to either the listen vs. control blocks or the test vs. listen blocks. The primary comparison of interest in both the GLM and ICA analysis was between the listen and control blocks (to determine areas more active during exposure to spoken language input than to acoustically related input that is not linguistic). A secondary interest was between the listen and test blocks. This comparison was intended to identify areas that were more active in either the listening or test blocks, given that both involved spoken language input, but only the test blocks involved decision making (responses to test items). The images related to the GLM analysis were thresholded at $p < .05$, with family-wise error (FWE) correction for brain-wide comparisons.

The ICA analysis also used the preprocessed data described above. We used the Extended Infomax algorithm in GIFT software (version 1.3i; Rachakonda, Egolf, Correa, & Calhoun, 2007). This approach has been shown to have a higher sensitivity to true signal change compared with other methods (Arya, Calhoun, Roys, Adali, Greenspan, & Gullapalli, 2003). Individual subject data were concatenated (Schmithorst & Holland, 2004) and the scan number (Scan 1-4) was specified as a repeated measure so that the individual ICs produced for each scan were identified in an identical manner across sequential scans (e.g., the first IC in scan 1 was also identified as the first IC in scans 2-4). The product of this analysis is a set of waveforms that represents signal variation that is independent of other sources of signal variation. GIFT maximizes spatial independence among the ICs produced by the ICA. This tends to produce ICs that are more stable in their location than when temporal independence of ICs is maximized (Calhoun, Eichele, Adali, & Allen, 2012).

We used GIFT to estimate the optimal number of ICs to be calculated so that the balance between under-fitting and over-fitting the data was optimized. The algorithm and its effectiveness are described in Li, Adali, & Calhoun, (2007). We allowed for 22 components to be generated during the analysis. We used the full set of subjects to estimate this number of components because this approach yields the best possible estimate. We ran 10 iterations of the ICA analysis. This is needed because ICA is a stochastic statistic, which can result in

slight variations in the ICs produced over multiple calculations (Himberg, Hyvärinen, & Esposito, 2004). Multiple iterations permit assessment of the stability of each IC component, and stability estimates can be quantitatively evaluated (see below).

We expected that 22 ICs would exceed the number of task-related components. Recall that ICA will also model nuisance variables like residual subject movement and physiologic noise. Likewise, a number of ICs were likely to account for little signal variance or to be idiosyncratic relative to the periods of stimulus exposure. To eliminate such ICs from further consideration, we performed a sequence of steps to differentiate between task-related and non-task related ICs. First, the stability of the IC components was evaluated statistically using the Icasto option in GIFT. The first twenty ICs had Iq values of .91 or above, indicating acceptable stability (maximum possible value = 1.00). We further considered only ICs for which the group activation t-values both positive and significant for either the listening blocks compared with the control blocks (listen > auditory stimulation) or the test block compared with the listening block (listening+decision > listening). There were eight ICs that met this criterion. The corresponding activation maps for these ICs were then thresholded, based on group t-statistics, at a corrected level of $p < .05$ (FWE correction). These maps represented the spatial distribution of signal associated with each IC waveform for the group as a whole. The eight maps were visually inspected for signal artifact by two examiners. These examiners looked for characteristic signatures of movement (e.g., whole head movement, fluid movement) and susceptibility artifact. Four of the eight ICs contained signal that was judged to primarily reflect artifact-related signal. This was expected, given that with ICA, these sources of variance associated with artifacts are segregated from task-related activation because they each tend to have their own unique temporal signature. This is true even when motion correction has been applied during preprocessing (McKeown, Hansen, & Sejnowski, 2003).

Eliminating artifact-related ICs left four ICs that were associated with either the listening phase of the scan, the test phase, or both, at a p value of .05 or less for at least one of the four scans. Requiring a component to reach this criterion during at least one scan, rather than during all four scans, allowed us to look at ICs that may be weak during a particular phase of the experiment, but strong at another point. This statistical criterion yielded four components that were associated with the listening vs. control block period, and three of these four were stronger for test vs. listening. There were no ICs that were active only for the test vs. listening comparison. These four components were highly stable, with Iq values of over .97 each. Using back reconstruction, we extracted the signal change associated with each IC, on a subject-by-subject basis for each of the four scans. This procedure allowed us to obtain data on activation strength for IC in each of the four scans.

For descriptive purposes, we also calculated laterality indices based on signal change for the IC components using the formula $L - R / ((L + R) * .5)$. Thus, positive laterality indices indicate left lateralization of the IC and negative indices indicate right lateralization.

Individual ICs can include multiple foci of activation, each of which maybe differentially activated across the four scans. To further describe the distribution and regional strength of activation within each IC, we identified activation peaks within statistically-significant areas

identified for each of the four ICs. Because this information is not the primary focus of our analyses, this is presented in Supplemental Tables 1-4 and are intended to permit more focused regional comparisons between scans or between studies.

Activation for each IC connected a number of anatomically-diverse regions (see Figure 2). In order to describe regional activation within each IC, we needed to separate peaks of activation within each IC. To do this, we eroded the already-thresholded images by 5%. This reduced the amount of diffuse activation that conjoined separate peaks of activation within the IC. Visual inspection indicated that this separated a majority, but not all activation peaks within the IC. We further separated regions containing two or more activation peaks only when the peaks were separated by an anatomical boundary (e.g., inferior temporal lobe and superior cerebellum, superior parietal lobule and supramarginal gyrus). When adjacent peaks did not cross anatomical boundaries (e.g., two adjacent peaks within the premotor strip), these were not separated further. The maximum *t* value was then extracted for each region. We used the probabilistic cortical, subcortical, and cerebellar atlases in FSLView version 3.1 (FSL, Jenkinson, Beckman, Behrens, Woolrich, & Smith, 2012) to identify the XYZ coordinates for each activation peak.

3. Results

3.1 Behavioral Results

There was a missing data point for one subject in Scan 1, due to equipment failure. For the remaining data, *d* prime was calculated to reflect the participants' ability to distinguish between correct and incorrect sentences during the test blocks of each scan. These data are presented in Figure 1a. The *d* prime statistic for scan 1 was not significantly above chance, despite the fact that participants had heard the stimuli both just prior to this scan and during this scan (approximately 4 minutes of exposure). However, behavioral performance was significantly above chance ($p < .05$, corrected for multiple comparisons) after scan 2 (approximately 6 and a half minutes of exposure) and remained so for the remaining scans. Although Figure 1a suggests a trend towards improved performance across the four scans, this was not statistically significant ($F(3,36) = .599$, $p = .619$).

The participants' subjective ratings of task difficulty are displayed in Figure 1b. As this graph demonstrates, participants rated the task as relatively difficult throughout the four scans, with ratings hovering around 4 on the 5 point scale for which a rating of 5 indicated the task was extremely difficult. There were no significant differences between scans in this behavioral rating (Friedman ANOVA, Chi square=1.894, $p = .594$).

3.2 fMRI Analysis

The four ICs that met criteria for stability and task-relatedness are displayed in Figure 2. We also provide the GLM results in Supplemental Figure 1 for comparison purposes. As expected from the nature of the analysis, the ICA results are similar to the GLM results, with the former returning more extensive areas of significant activation than the latter. As Figure 2 indicates, the regional activation that characterized each IC was relatively similar across scans, with more overlap than separation across time. Furthermore, there was more

separation than overlap between the spatial distribution of different ICs. This is a direct result of mathematically maximizing the spatial independence of the components during ICA derivation.

The ICA identified four sets of functionally-connected regions that were task-related. The ideal waveform for each of the four ICs showed significant positive correlations with the listening vs. control blocks of the fMRI scan and IC-1, IC-2, and IC-4 also showed a significant positive correlation with the test vs. listening blocks. The lateralization index indicated that IC-2 was consistently left lateralized for all four scans, with laterality values ranging from 0.33 to 0.74 for the group as a whole. IC-3 was consistently right lateralized for all four scans, with values ranging between -0.30 and -0.58. The remaining two ICs did not show a consistent laterality across scans and laterality values were relatively low, ranging from -0.17 to 0.19 for IC-1 and -0.20 to 0.002 for IC-4.

The percent signal change associated with each IC showed unique patterns of activation across the four scans. These are displayed as graphs in Figure 2. In these graphs, we split the participant group into low and high performance groups based on a median split of behavioral performance during that scan. Recall that performance was below chance during scan 1. In scan 2, the low performing group were those with d' values below 0.46 (mean=-0.03) and the high performing group were above this value (mean=0.85). For scan 2, the low performing group included d' scores below 0.45 with a mean of 0.28 and the mean for the high performing group was 0.80. For scan 3, the low performing group included d' values below 0.59 with a mean of 0.40, and the mean for the high performing group was 0.86. For scan 4, the low performing group included d' values below 0.72 with a mean of 0.35 and the mean for the high performing group was 1.0. Note that by scan 2, when behavioral performance was above chance for the group as a whole, there was a general pattern for greater activation by the low performers relative to the high performers.

3.3 Brain-behavior relations

We asked whether behavioral performance (d prime) was correlated with activation in one or more of the ICs. We first divided each IC into right and left hemisphere regions. This was done because two of the four ICs (ICs 2 and 3) displayed consistently lateralized activation throughout the experiment. This raised the possibility that the brain-behavior relation supported by these ICs would also show a greater effect for one hemisphere relative to the other. IC-1 correlated significantly with behavioral performance in scan 2 and this correlation was significant for both left ($r=-.57$) and right hemisphere activation ($r=-.66$). For IC-2, a left lateralized component, only left hemisphere activation correlated with behavior ($r=-.57$), and only during scan 2. For the right lateralized IC-3, right hemisphere activation correlated with behavior during scan 2 ($r=-.55$) and scan 3 ($r=-.58$). In addition, left hemisphere activation in this IC correlated with behavior only during scan 3 ($r=-.54$). IC4, which was not consistently lateralized across scans, correlated with behavior in the left hemisphere during scan 3 ($r=-.56$). Note that in every case, significant correlations occurred during scans when behavioral performance for the group was above chance levels. There were no significant correlations between an IC and performance during scan 1. Performance was below chance during this scan.

4. Discussion

Initial processing of an unfamiliar natural language was associated with four task-related ICs. These components as a group included regions previously reported in imaging studies of artificial languages. These included the posterior temporal and superior temporal regions that have consistently been reported for word segmentation studies (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2006) as well as the inferior frontal activation reported for finite state grammar learning (Newman-Norlund et al., 2006). In the present study, these anterior and posterior regions activated on different IC time courses, suggesting independent contributions to the overall language-learning network. We also observed additional regions of significant activation that encompassed frontal and parietal/lateral occipital activations. Although middle frontal activations (both left hemisphere and right hemisphere) have been reported in artificial language studies (Cunillera et al., 2009; McNealy et al., 2006; Newman-Norlund et al., 2006), the more posterior activations have not been reported previously. This latter difference is unlikely to be due to the differences between the GLM analysis used in previous studies and the ICA used as the primary analysis here. Parietal activation was detected by our GLM analyses as well as by ICA.

The pattern of activation across the four ICs and their correlations with behavioral performance suggests differential contribution to each IC at different points during the period of exposure to the Icelandic stimuli. During scan 1, no IC correlated with behavioral performance, and performance did not differ from chance. Behavioral performance was above chance by scan 2 and behavior correlated with a subset of ICs in scans 2 and 3. In each case, correlations were negative, suggesting that greater physiological effort was expended by those who were the least successful during that scan. During scan 4, behavioral performance remained above chance, but there were no correlations with any of the ICs. It is possible that participants were habituating to now-familiar stimuli, accounting for the lack of correlation with behavior for the final scan. Alternately, scan 4 may reflect a transition point in the learning process and additional ICs or activation peaks may have emerged had the experiment continued over a longer period.

The spatial distribution of the four ICs, as well as their pattern of signal change, indicate differential contributions to processing over time. IC1 contains regions of activation most similar to those reported for word segmentation tasks in previous artificial language studies (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2006), which have consistently included bilateral temporal activations and often include lateral prefrontal or middle frontal activations. Although this temporal region is often linked to semantic processing, this role can be ruled out in the present experiment because participants did not have access to the meaning of the Icelandic sentences they heard. However, the mid-superior temporal sulcus also responds to non-meaningful acoustic information (Altmann, Bledowski, Wibrall, & Kaiser, 2007; Altmann, Henning, Döring, & Kaiser, 2008; Specht, Osnes, & Hugdahl, 2009; Sturm, Schnitker, Grande, Huber, & Willmes, 2011) as well as phoneme-level information (Turkeltaub & Coslett, 2010; Sturm et al., 2011). Particularly relevant to the present study, this region activates specifically if listeners expect to hear speech when given ambiguous auditory stimuli (Osnes, Hugdahl, Hjelmervik, & Specht, 2012) and this region responds incrementally as listeners begin to recognize phonetic content in an auditory signal (Specht

et al., 2009). Furthermore, recognition of foreign-language phoneme contrasts is correlated with activation in this region (Callan, Jones, Callan, & Akahane-Yamada, 2004). The temporal lobe activations in these studies are consistent with the temporal activations of IC 1. This IC also contains activation in the premotor cortex, which has been previously reported in an artificial language learning study (Karuza et al., 2013). This region has been linked to speech perception (Wilson, Saygin, Sereno, & Iacoboni, 2004) and is thought to reflect engagement of motor cortex during speech processing (Hickok, 2012). Given prior reports of functional connectivity between the premotor and superior temporal sulcus region (Chevillet, Jiang, Rauschecker, & Riesenhuber, 2013; Osnes et al., 2011; Wison & Iacoboni, 2006), it is not surprising that these two areas are found within the same independent component here.

Like IC-1, left IC-2 and right IC-3 were also negatively correlated with behavior in scan 2. IC-3 correlated (bilaterally) with behavioral performance during scan 3 as well. The frontal-parietal/occipital distribution of these two ICs strongly suggests a role in executive control of the processing of the input. Similar activations are widely reported during basic auditory processing tasks (e.g., Anderson, Ferguson, Lopez-Larson, & Yurgelun-Todd, 2010; de Souza, Yehia, Sato, & Callan, 2013; Henry, Herrmann, & Obleser, in press; Sabri, Humphries, Binder, & Liebenthal, 2008; Shaywitz, Shaywitz, Pugh, Fulbright, Skudlarski, Mencl, & Gore, 2001; Sturm et al., 2011) and these effects are typically bilateral. In the present study, bilateral activation was seen, but it was segregated across two independent components, which showed left and right hemisphere lateralization respectively. This segregation suggests that the bilateral activity previously reported may reflect two different functional contributions differentially represented by the two hemispheres.

The fact that behavior was negatively correlated with these ICs, and activation was relatively low for these ICs in scan 2 is consistent with the notion that down-regulation of frontal cortex may facilitate some kinds of processing (Chrysikou, Hamilton, Coslett, Datta, Bikson, & Thompson-Schill, 2013; Chrysikou & Thompson-Schill, 2011; Thompson-Schill & Ramscar & Chrysikou, 2009). In particular, inhibitory stimulation to the left frontal lobe resulted in faster reaction times on a task that required inhibiting a prototypical response in favor of a novel response (Chrysikou et al., 2013). Although this particular study involved visual input, it offers a potential parallel to novel language learning in which learners may need to inhibit their native language biases in order to learn the patterns that represent a new language.

Left lateralized activations, similar to that in IC-2, have also been reported for improvements in processing of nonspeech sound contrasts (de Souza et al., 2013) and under conditions that vary attentional load for processing non-meaningful consonant-vowel stimuli (Westerhausen, Moosmann, Alho, Belsby, Hämäläinen, Medvedev, & Hugdahl, 2010). For this reason, we suggest that IC-2 may have reflected attention to short-duration acoustic information in the speech stream. This IC also contained an activation peak in the posterior left inferior temporal and fusiform gyri, which first shows suprathreshold activation (significant *t* values) during scan 2. This general area, which is frequently associated with visual word-form recognition, is also known to activate when attention is focused on auditorily-presented words (Yoncheva, Zevin, Maurer, & McCandliss, 2009). Therefore, this

IC may reflect the application of attentional resources to sublexical and lexical aspects of the Icelandic words.

The right lateralized nature of IC-3 suggests that the focus of attention was not the short-duration acoustic information that has characterized the left-lateralized activity of these areas in IC-2. The right lateralized fronto-parietal-temporal distribution is consistent with that found when listeners are asked to discriminate sentences based on their prosodic envelopes (Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, et al., 2003; Plante, Creusere, & Sabin, 2002) and the right hemisphere activation is enhanced for listeners processing a foreign language (Gandour et al., 2003; Gandour, Tong, Wong, Talavage, Dzemidzic, Xu, & Lowe, 2004). The frontal and parietal activations are often absent when attention is not specifically focused on evaluating the prosodic information by the task demands (Plante et al., 2002; Humphreys, Love, Swinney, & Hickok, 2005; Ischebeck, Friederici, & Alter, 2008), suggesting that the primary activations in IC-3 are related to attention to prosodic information in the input.

IC-4, like IC-1, included posterior temporal and middle frontal activation. However, these two ICs, which by definition activated on independent time courses, also showed different activation strengths and different correlations with behavior over time. Furthermore, IC-4 contained a prominent activation in the regions of the inferior frontal gyrus and the anterior cingulate/supplemental motor cortex that was not found in IC-1. These differences strongly suggests that the two ICs were engaged in processing different aspects of the auditory input. The activation in inferior frontal gyrus was unique to IC-4 and has been reported in other artificial language studies (Karuza et al., 2013; Newman-Norlund et al., 2006). Others have suggested that the left inferior frontal gyrus in particular is involved in the accumulation of information during processing of ongoing information (Nopenny, Ostwald, & Werner, 2010) including identification of individual words in speech segmentation tasks (Karuza et al., 2013) and detection of legal word sequences representing a finite state grammar (Newman-Norlund et al., 2006).

There is an alternate explanation for the role of IC-4, which is not mutually exclusive with the idea that it is contributing to word-level statistics. It is not clear in the present study, or in previous artificial language studies (Karuza et al., 2013; Newman-Norlund et al., 2006), whether the inferior frontal effect can be attributed to accumulation of statistical information or memory for units already segmented from the speech stream (or both). In the present study, this IC was correlated both with the listening and test phases of the scan, suggesting processes common to learning and test phases, where item memory could be employed. Similarly, activity in the inferior frontal gyrus was associated with test performance in Karuza et al. (2013), but was not detected by an omnibus test of activation associated with simply listening to the stimuli. The correlation with post-scan behavioral performance suggests a role in memory encoding. This idea is supported by reports that activation within the inferior frontal gyrus and areas in the temporal-parietal junction during memory encoding tasks predict later recognition of verbal stimuli (e.g., Casasanto, Killgore, Maldjian, Glosser, Alsop, Cooke, Grossman, & Detre, 2002; Clark & Wagner, 2003; Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006).

In summary, the results support the hypothesis that dynamic changes in regional activation occur within the initial minutes of exposure to an unfamiliar natural language. The ICA analysis identified spatially-distinct regions of activation that showed different patterns of signal change across four consecutive scans and that showed differential relations to behavioral performance over time. The behavioral data further indicate that, although learning did occur, our participants had not fully acquired all aspects of the language that were reflected in the input. Thus it is not surprising that the ICA did not indicate a steady processing state had been achieved before the end of the experiment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was partially funded by NIH grants 1R01 DC011276 and DC-01409.

Appendix

Block	Icelandic Stimuli	English Translation
Listen	Kaldi hvíti fiskurinn synti. Gamli maðurinn átti bát Lítill drengur hoppaði hratt Kaldur maður synti varlega Gamall báturinn sigldi hratt á vatninu Maður sigldi bát Gamall bátur sigldi varlega	The cold white fish swam The old man owned a boat The little boy jumped fast The cold man swam carefully An old boat sailed fast on the lake A man sailed a boat An old boat sailed carefully
Test	Drengurinn hoppaði hratt Maðurinn synti varlega *Hratt einn sigldi bátur *Synti fiskur kaldur *Rak gamall einn Hvítur fiskur synti	The boy jumped fast The man swam carefully *Fast one sailed boat *Swam fish cold *Drifted old one A White fish swam
Listen	Einn gamall bátur sigldi hratt Kaldi hvíti fiskurinn sá gamla manninn Einn gamall fiskur hoppaði Gamli maðurinn sá stóra manninn Kaldur drengur hoppaði hátt á bátinum Drengurinn hoppaði hátt Kaldi hvíti fiskurinn synti	One old boat sailed fast The cold old fish saw the old man One old fish jumped The old man saw the big man The cold boy jumped high on the boat The boy jumped high The cold white fish swam
Test	*Hvítur gamall rak Gamall maður synti hratt *Hvítur drengur vatnið Gamall maður sá dreng *Hvítur bát fiskur Drengurinn rak hratt	*White old drifted An old man swam fast * White boy the water An old man saw a boy *White boy fish The boy drifted fast
Listen	Drengurinn sá gamlann mann í bátinum Drengurinn hoppaði hátt Gamli maðurinn átti bát Báturinn rak Gamall maður var með dreng Hvítur maður sigldi hratt	The boy saw an old man in the boat The boy jumped high The old man owned a boat The boat drifted An old man was with a boy A white man sailed fast
Test	*Drengur sá mann hoppaði Gamall maður var með gömlum dreng *Vatn sá hoppaði hratt Gamlann mann sá drengurinn Einn gamall drengur var á bát *Hoppaði drengur	*Boy saw man jump Old man was with an old boat *Lake saw jumped fast An old man saw the boy One old boy was on a boat *Jumped boy

Block	Icelandic Stimuli	English Translation
Listen	Dregurinn var með gamla manninum Bátur sigldi varlega Dregurinn sigldi bátnum Báturinn sigldi niður ána Dregurinn sá gamla manninn á bátnum Dregurinn hoppaði á vatninu Gamall maður var með dreng Dregurinn sá gamla manninn á bátnum	The boy was with the old man A boat sailed carefully The boy sailed the boat The boat sailed down the river The boy saw the old man on the boat The boy jumped on the lake An old man was with a boy The boy saw the old man on the boat
Test	*Maður sigldi kaldur Dregur sigldi bát Maðurinn hoppaði á bátnum *Bátur hoppaði sigldi hoppaði Dregur var með gömlum manni *Ekki var maður gamli	*A man sailed cold A boy sailed a boat The man jumped on the boat *A boat jumped sailed jumped A boy was with an old man *Not was man old
Listen	Svartur drengur sigldi hratt Báturinn rak niður ána Hvíti drengurinn hoppaði Gamli hvíti báturinn rak á ánni Ungi maðurinn var með gamla manninum Báturinn rak hratt á ánni Einn gamall maður sigldi Hvítur maður hoppaði Dregurinn sigldi bátnum	A black boy sailed fast The boat drifted down the river The white boy jumped The old white boat drifted on the river The young man was with the old man The boat drifted fast on the river One old man sailed A white man jumped The boy sailed the boat
Test	Gamall hvítur fiskur hoppaði *Gamall bátur ungur drengur Gamli maðurinn sá drenginn *Maður gamall strákur *Hoppaði drengur kaldur *Maður bátur rak varlega	An old white fish jumped *Old boat young boy The old man saw the boy *The man old boy *Jumped boy cold *Man boat drifted carefully
Listen	Hvíti drengurinn sá vatnið Maðurinn sigldi bátnum hratt Dregurinn sigldi bátnum Dregurinn var á bátnum Kaldur drengur sá gamla manninn Dregurinn sá gamla manninn Dregurinn synti á vatninu Gamli maðurinn sigldi gamla hvíta bátnum	The white boy saw the water The man sailed the boat fast The boy sailed the boat The boy was on the boat The cold boy saw the old man The boy saw the old man The boy swam on the lake The old man sailed the old white boat
Test	Dregur hoppaði á bát Dregurinn sigldi bátnum Hvíti drengurinn sigldi gamla hvíta bátnum Dregur hoppaði á bát Sigldi hvítur bátur gamall Dregur hoppaði gamall maður	A boy jumped on a boat The boy sailed the boat The white boy sailed the old white boat A boy jumped on a boat *Sailed white boat old *Boy jumped old man

References

- Altmann C, Bledowski C, Wibrall M, Kaiser J. Processing of location and pattern changes of natural sounds in the human auditory cortex. *Neuroimage*. 2007; 35:1192–1200. [PubMed: 17320413]
- Altmann C, Henning M, Döring M, Kaiser J. Effects of feature-selective attention on auditory pattern and location processing. *Neuroimage*. 2008; 41:69–79. [PubMed: 18378168]
- Anderson J, Ferguson M, Lopez-Larson M, Yurgelun-Todd D. Topographic maps of multisensory attention. *Proceedings Of The National Academy Of Sciences Of The United States Of America*. 2010; 107:20110–20114. [PubMed: 21041658]
- Arya R, Calhoun VD, Roys S, Adali T, Greenspan J, Gullapalli R. Comparative study of several multivariate fMRI processing methods: PCA, Factor Analysis, Infomax, FASTICA, MELODIC. Presented at the Proceedings of the International Society for Magnetic Resonance in Medicine. 2003:1–1.
- Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*. 1998; 9:321–324.

- Callan D, Jones J, Callan A, Akahane-Yamada R. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage*. 2004; 22:1182–1194. [PubMed: 15219590]
- Cappa S. Imaging semantics and syntax. *Neuroimage*. 2012; 61:427–431. [PubMed: 22019859]
- Casasanto DJ, Killgore WDS, Maldjian JA, Glosser G, Alsop DC, Cooke AM, Grossman M, Detre JA. Neural correlates of successful and unsuccessful verbal memory encoding. *Brain and Language*. 2002; 80:287–295. [PubMed: 11896642]
- Chevillet M, Jiang X, Rauschecker J, Riesenhuber M. Automatic phoneme category selectivity in the dorsal auditory stream. *The Journal Of Neuroscience*. 2013; 33:5208–5215. [PubMed: 23516286]
- Chrysikou E, Hamilton R, Coslett H, Datta A, Bikson M, Thompson-Schill S. Noninvasive transcranial direct current stimulation over the left prefrontal cortex facilitates cognitive flexibility in tool use. *Cognitive Neuroscience*. 2013; 4:81–89. [PubMed: 23894253]
- Clark D, Wagner AD. Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. *Neuropsychologia*. 2003; 41:304–317. [PubMed: 12457756]
- Chrysikou E, Thompson-Schill S. Dissociable brain states linked to common and creative object use. *Human Brain Mapping*. 2011; 32:665–675. [PubMed: 20533561]
- Cunillera T, Càmarà E, Toro J, Marco-Pallares J, Sebastián-Galles N, Ortiz H, Rodríguez-Fornells A. Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*. 2009; 48:541–553. [PubMed: 19580874]
- De Diego Balaguer R, Toro J, Rodríguez-Fornells A, Bachoud-Lévi A. Different neurophysiological mechanisms underlying word and rule extraction from speech. *Plos One*. 2007; 2:e1175. [PubMed: 18000546]
- de Souza A, Yehia H, Sato M, Callan D. Brain activity underlying auditory perceptual learning during short period training: simultaneous fMRI and EEG recording. *BMC Neuroscience*. 2013; 148
- Fletcher P, Büchel C, Josephs O, Friston K, Dolan R. Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cerebral Cortex*. 1999; 9:168–178. [PubMed: 10220229]
- Gandour J, Dziedzic M, Wong D, Lowe M, Tong Y, Hsieh L, et al. Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*. 2003; 84:318–336. [PubMed: 12662974]
- Gandour J, Tong Y, Wong D, Talavage T, Dziedzic M, Xu Y, Lowe M. Hemispheric roles in the perception of speech prosody. *Neuroimage*. 2004; 23:344–357. [PubMed: 15325382]
- Glover GH, Law CS. Spiral-In/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magnetic Resonance in Medicine*. 2001; 46:515–522. [PubMed: 11550244]
- Henry M, Herrmann B, Obleser J. Selective Attention to Temporal Features on Nested Time Scales. *Cerebral Cortex*. (in press).
- Hickok G. The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *Journal Of Communication Disorders*. 2012; 45(6):393–402. [PubMed: 22766458]
- Himberg J, Hyvärinen A, Esposito F. Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage*. 2004; 22:1214–1222. [PubMed: 15219593]
- Humphries C, Love T, Swinney D, Hickok G. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*. 2005; 26:128–138. [PubMed: 15895428]
- Iidaka T, Matsumoto A, Nogawa J, Yamamoto Y, Sadato N. Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analysis using fMRI and ERP. *Cerebral Cortex*. 2006; 16:1349–1360. [PubMed: 16861334]
- Ischebeck A, Friederici A, Alter K. Processing prosodic boundaries in natural and hummed speech: an FMRI study. *Cerebral Cortex*. 2008; 18:541–552. [PubMed: 17591598]
- Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012; 62:782–790.

- Jusczyk P. Narrowing the distance to language: one step at a time. *Journal Of Communication Disorders*. 1999; 32(4):207–222. [PubMed: 10466094]
- Karunanayaka P, Schmithorst VJ, Vannest J, Szaflarski JP, Plante E, Holland SK. A Group Independent Component Analysis of Covert Verb Generation in Children: A Functional Magnetic Resonance Imaging Study. *NeuroImage*. 2010; 51:472–487. [PubMed: 20056150]
- Karuzs E, Newport E, Aslin R, Starling S, Tivarus M, Bavelier D. The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain And Language*. 2013; 127:46–54. [PubMed: 23312790]
- McKeown MJ, Hansen LK, Sejnowski TJ. Independent component analysis of functional MRI: what is signal and what is noise? *Current opinion in Neurobiology*. 2003; 13:620–629. [PubMed: 14630228]
- McNealy K, Mazziotta J, Dapretto M. Cracking the language code: neural mechanisms underlying speech parsing. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*. 2006; 26:7629–7639. [PubMed: 16855090]
- McNealy K, Mazziotta J, Dapretto M. The neural basis of speech parsing in children and adults. *Developmental Science*. 2010; 13:385–406. [PubMed: 20136936]
- Newman-Norlund R, Frey S, Petitto L, Grafton S. Anatomical substrates of visual and auditory miniature second-language learning. *Journal Of Cognitive Neuroscience*. 2006; 18:1984–1997. [PubMed: 17129186]
- Noppeney U, Ostwald D, Werner S. Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*. 2010; 30(21):7434–7446. [PubMed: 20505110]
- Osnes B, Hugdahl K, Specht K. Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage*. 2011; 54:2437–2445. [PubMed: 20932914]
- Opitz B, Ferdinand N, Mecklinger A. Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Frontiers In Human Neuroscience*. 2011; 58:1–9.
- Osnes B, Hugdahl K, Hjelmervik H, Specht K. Stimulus expectancy modulates inferior frontal gyrus and premotor cortex activity in auditory perception. *Brain And Language*. 2012; 121:65–69. [PubMed: 22377261]
- Peña M, Bonatti LL, Nespor M, Mehler J. Signal-driven computations in speech processing. *Science*. 2002; 298:604–607. [PubMed: 12202684]
- Plante E, Creusere M, Sabin C. Dissociating sentential prosody from sentence processing: Activation interacts with task demands. *NeuroImage*. 2002; 17:401–410. [PubMed: 12482093]
- Plante E, Schmithorst VJ, Holland SK, Byars AW. Sex Differences in the Activation of Language Cortex During Childhood. *Neuropsychologia*. 2006; 44:1210–1221. [PubMed: 16303148]
- Poldrack R, Rodriguez P. How do memory systems interact? Evidence from human classification learning. *Neurobiology Of Learning And Memory*. 2004; 82:324–332. [PubMed: 15464413]
- Price C. The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals Of The New York Academy Of Sciences*. 2010; 1191:62–88. [PubMed: 20392276]
- Price C. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 2012; 62:816–847. [PubMed: 22584224]
- Rachakonda S, Egolf E, Correa N, Calhoun V. Group ICA of fMRI Toolbox (GIFT) Manual. 2007 Available at <http://mialab.mrn.org/software/#gica> (Downloaded December 14, 2012).
- Romberg AR, Saffran JR. All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*. 2013; 37:1290–1318. [PubMed: 23772795]
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science*. 1996; 274:1926–1928. [PubMed: 8943209]
- Saffran JR, Aslin RN, Newport EL. Word segmentation: The role of distributional cues. *Journal of Memory and Language*. 1996; 35:606–621.
- Saffran JR, Aslin RN, Newport EL, Tunick RA, Barrueco S. Incidental language learning: Listening (and Learning) out of the corner of your ear. *Psychological Science*. 1997; 8:101–105.

- Sabri M, Humphries C, Binder J, Liebenthal E. Neural events leading to and associated with detection of sounds under high processing load. *Human Brain Mapping*. 2013; 34(3):587–597. [PubMed: 22102362]
- Schmithorst VJ, Holland SK. Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data. *Journal of Magnetic Resonance Imaging*. 2004; 19:365–368. [PubMed: 14994306]
- Schmithorst VJ, Holland SK, Plante E. Cognitive Modules Utilized for Narrative Comprehension in Children: A Functional Magnetic Resonance Imaging Study. *NeuroImage*. 2006; 29:254–266. [PubMed: 16109491]
- Schmithorst VJ, Holland SK, Plante E. Development of effective connectivity for narrative comprehension in children. *NeuroReport*. 2007; 18:1411–1415. [PubMed: 17712265]
- Schmithorst VJ, Holland SK, Plante E. Object Identification and Lexical/Semantic Access in Children: A Functional Magnetic Resonance Imaging Study of Word-Picture Matching. *Human Brain Mapping*. 2007; 28:1060–1074. [PubMed: 17133401]
- Shaywitz B, Shaywitz S, Pugh K, Fulbright R, Skudlarski P, Mencl W, Gore J. The functional neural architecture of components of attention in language-processing tasks. *Neuroimage*. 2001; 13:601–612. [PubMed: 11305889]
- Sony Pictures Digital, Inc. Sound Forge 7.0. Madison, WI: Author; 2003.
- Specht K, Osnes B, Hugdahl K. Detection of differential speech-specific processes in the temporal lobe using fMRI and a dynamic "sound morphing" technique. *Human Brain Mapping*. 2009; 30:3436–3444. [PubMed: 19347876]
- Sturm W, Schnitker R, Grande M, Huber W, Willmes K. Common networks for selective auditory attention for sounds and words? An fMRI study with implications for attention rehabilitation. *Restorative Neurology And Neuroscience*. 2011; 29:73–83. [PubMed: 21701059]
- Szaflarski JP, Altaye M, Rajagopal A, Eaton K, Meng XX, Plante E, Holland SK. A 10-year longitudinal fMRI study of narrative comprehension in children and adolescents. *NeuroImage*. 2012; 63:1188–1195. [PubMed: 22951258]
- Szaflarski JP, Schmithorst VJ, Altaye M, Byars AW, Rett J, Plante E, Holland SK. FMRI study of longitudinal language development in children age 5-1. *Annals of Neurology*. 2006; 59:796–807. [PubMed: 16498622]
- Thiessen ED, Saffran JR. When cues collide: Use of stress and statistical cues to word boundaries by 7- and 9-month old infants. *Developmental Psychology*. 2003; 39:706–716. [PubMed: 12859124]
- Thompson-Schill S, Ramscar M, Chrysikou E. Cognition without control: When a little frontal lobe goes a long way. *Current Directions In Psychological Science*. 2009; 18:259–263. [PubMed: 20401341]
- Turkeltaub P, Coslett H. Localization of sublexical speech perception components. *Brain And Language*. 2010; 114(1):1–15. [PubMed: 20413149]
- Veroude K, Norris D, Shumskaya E, Gullberg M, Indefrey P. Functional connectivity between brain regions involved in learning words of a new language. *Brain And Language*. 2010; 113:21–27. [PubMed: 20116090]
- Westerhausen R, Moosmann M, Alho K, Belsby S, Hämäläinen H, Medvedev S, Hugdahl K. Identification of attention and cognitive control networks in a parametric auditory fMRI study. *Neuropsychologia*. 2010; 48:2075–2081. [PubMed: 20363236]
- Wilson S, Iacoboni M. Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *Neuroimage*. 2006; 33(1):316–325. [PubMed: 16919478]
- Wilson S, Saygin A, Sereno M, Iacoboni M. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*. 2004; 7:701–702.
- Yoncheva Y, Zevin J, Maurer U, McCandliss B. Auditory selective attention to speech modulates activity in the visual word form area. *Cerebral Cortex*. 2010; 20:622–632. [PubMed: 19571269]

Highlights

- Participants were asked to learn an unfamiliar language over 4 consecutive fMRI scans
- Early-stage learning involves the dynamic application of resources over time.
- Different independent components correlate with behavioral performance over time.
- Natural language learning differs in part from prior artificial language learning studies.

Figure 1.

a. Behavioral performance (d' prime) on the test blocks for each of four fMRI scans. * indicates d' distributions differed significantly from chance. b. Participants' ratings of task difficulty on a five point scale where 0 is extremely easy and 5 is extremely difficult. Error bars indicate the standard error.

Figure 2.

Task-related independent components. The brain images indicate the spatial distribution of each independent component (thresholded at a corrected level of $p < .05$, FWE). Colors labeled 1 through 4 indicate IC spatial distributions unique to scans 1 through 4. Colors labeled +2 through +4 reflect overlap of the IC component among 2, 3, or 4 scans. The line graphs present the group average signal change for each component across the four scans, with bars indicating the standard error around the group mean. The group is subdivided into those who were low and high performers based on a median split of behavioral accuracy at each scan. * indicates scans in which the signal was significantly correlated with behavioral performance. Correlations were negative in all cases.