Published in final edited form as:

J Neurolinguistics. 2015 November 1; 36: 17–34. doi:10.1016/j.jneuroling.2015.04.005.

The nature of the language input affects brain activation during learning from a natural language

Elena Plante^a, Dianne Patterson^a, Rebecca Gómez^a, Kyle R. Almryde^{a,c}, Milo G. White^a, and Arve E. Asbjørnsen^b

^aThe University of Arizona Department of Speech, Language, & Hearing Sciences PO Box 210071, The University of Arizona, Tucson, AZ 85721-0071, USA

^bUniversity of Bergen Department of Biological and Medical Psychology University of Bergen Jonas Lies vei 91 5009 Bergen Norway

Abstract

Artificial language studies have demonstrated that learners are able to segment individual word-like units from running speech using the transitional probability information. However, this skill has rarely been examined in the context of natural languages, where stimulus parameters can be quite different. In this study, two groups of English-speaking learners were exposed to Norwegian sentences over the course of three fMRI scans. One group was provided with input in which transitional probabilities predicted the presence of target words in the sentences. This group quickly learned to identify the target words and fMRI data revealed an extensive and highly dynamic learning network. These results were markedly different from activation seen for a second group of participants. This group was provided with highly similar input that was modified so that word learning based on syllable co-occurrences was not possible. These participants showed a much more restricted network. The results demonstrate that the nature of the input strongly influenced the nature of the network that learners employ to learn the properties of words in a natural language.

Keywords

fMRI; language; transitional probabilities; statistical learning; second language acquisition

1. Introduction

When a language is unfamiliar to a listener, it is often not obvious initially where one word ends and another begins in continuous speech. Only a small percentage of words are uttered in isolation (about 9% based on reports by Brent & Siskind, [2001] and Fernald & Morikawa, [1993]). More typically, multiple words are heard as a nearly continuous speech

Corresponding author: E Plante, PhD, Phone: 011-520-621-5080, eplante@email.arizona.edu.

^cCurrently at the University of Chicago

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.

stream. The ability to detect individual words in running speech is a fundamental early requirement for language acquisition.

In the last two decades, experimental studies grounded in learning theory have suggest that the ability to segment words in running speech is a product of a more general tendency of listeners to track regularities (Gómez, 2006; Newport & Aslin, 2004; Saffran, 2003). In particular, a statistical learning framework proposes that learners track distributional information in their environment and use that information to extract structure and principles about the sensory input they receive. This learning is considered unguided, in that it is not necessary to focus learners on particular aspects of the input and learning occurs in the absence of overt feedback. This tendency to attend to distributional information is not limited to language input, but is general to many types of stimuli including visual symbol sequences (e.g., Reber, 1967; Turk-Browne, Jungé, & Scholl, 2005) and arrays (Aguilar & Plante, 2013; Fiser & Aslin, 2002;), environmental scenes (e.g., Brady & Oliva, 2008; Fiser & Aslin, 2002; Turk-Browne, Scholl, Johnson, & Chun, 2010) and tactile sequences (Conway & Christiansen, 2005). The ability to track statistical information allows learners to extract regularities that can be represented as conceptual units. In the case of language, these conceptual units can define word boundaries, as well as grammatical relations within and between words.

Early evidence supporting this general theory comes from studies of word segmentation by infant learners. Infants can segment words from an artificial language by tracking the transitional probability of syllables in running speech (e.g., Aslin, Saffran, & Newport, 1998; Graf Estes, Evans, Alibali, & Saffran, 2007; Saffran, Aslin, & Newport, 1996; Thiessen, 2010; Thiessen & Saffran, 2003). In these types of experiments, high transitional probabilities among sequential syllables predict the presence of multisyllabic words whereas low transitional probabilities reflect the boundaries between words. Transitional probabilities are a more refined statistic than simple co-occurrence frequency, in that the former takes the base frequency of individual syllables into account.

The ability to extract even a few words from running speech further assists the identification of additional words (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Cunillera, Càmarà, Laine, & Rodríguez-Fornells, 2010). Indeed, some have suggested that the ability to track transitional probabilities to segment words from running speech provides the learner with information about phonotactic constraints, which further serves the process of word identification (Adriaans & Kager, 2010). Because morphemes (including words) form the basis of higher order linguistic generalizations, the initial process of word segmentation is a critical first step to discovering the structure of a language.

Although statistical learning of word forms was first described in infants using artificial languages, similar studies have demonstrated that adults can also use transitional probabilities to segment words from running speech. Such learning appears to be age-invariant, with robust performance reported for older typically-developing children (Evans, Saffran, & Robe-Torres, 2009; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) and adult learners (e.g. Cunillera et al., 2009; De Diego Balaguer, Toro, Rodríguez-Fornells, &

Bachoud Lévi, 2007; Saffran, Newport, & Aslin 1996 ; Saffran et al., 1997; Thiessen, 2010) as well as infants.

The artificial languages typically used in statistical learning studies can be criticized because they are dissimilar to natural languages in several important respects. Artificial language paradigms have typically used strings of consonant-vowel (CV) triplet pseudo-words that recur with high frequency and high density within the artificial language. For example, Aslin et al., (1998) used a four-item nonword corpus where each nonword occurred 45-90 times. This is strikingly different from natural languages, in which most words are not repeated within a sentence and certainly repeat less frequently across sentences than they do in artificial languages. Furthermore, the CVCVCV nonwords used in previous artificial language studies reflect a much more restricted word form than occurs in natural languages. There is evidence that artificial languages in which word forms are more similar to those found in natural languages are more difficult to learn (Johnson & Tyler, 2010). Therefore, it is clear that natural languages present a greater challenge for statistical learning than do artificial languages.

Recently, researchers have begun to investigate reliance on sequential statistics in the context of natural languages. Pelucchi, Hay, and Saffran (2009) showed that English-learning 8 month olds can use transitional probabilities to segment target words in Italian speech. Kittleston, Aguilar, Tokerud, Plante, & Asbjørnsen (2010) showed that adults with no knowledge of Norwegian can use transitional probabilities to segment target words in Norwegian speech. The stimuli used both by Pelucchi et al. and Kittleston et al. were more complex than those used in typical artificial language studies. Others have used a natural language (Mandarin) presented in the context of visual information (a video clip) that contained multiple cues to word units and their meanings (Gullberg, Roberts, Dimroth, Veroude, & Indefrey, 2010). These studies indicate that learners can use co-occurrence information, with or without additional visual cues, to segment words from running speech in a variety of natural languages.

These natural language studies lend ecological validity to the earlier behavioral work on word segmentation from artificial language paradigms. However, it is important to note that natural languages offer a variety of cues to the properties of words, any of which can assist the learner in segmenting words from running speech. Indeed, sequential statistical information characterizes only a portion of the words in speech that infants need to segment (Yang, 2004). Knowing the predominant stress pattern of one's native language is also thought to be a particularly useful word segmentation cue (Cutler & Norris, 1988; Jusczyk, Cutler, & Redanz, 1993). However, the predominant stress pattern for words in a language cannot be inferred before the learner has heard a sufficient corpus over which this pattern can be detected. This may be why infants show an early reliance on transitional probabilities, with reliance later in development on predominant stress as a cue to word boundaries (Thiessen & Saffran, 2003). Likewise, phonotactic information offers another cue to word structure. For example, English words very rarely begin with /ts/, but this sound sequence is reasonably common in word final position. Phonotactics can be used as a cue for word segmentation, but requires sufficient exposure to a language for learners to use these patterns (Adriaans & Kager, 2010). Therefore, it is not possible to control for all sources of

statistical information in natural language input. However, it is possible to manipulate the input so that learners do or do not have access to the transitional statistics that facilitate rapid identification of word boundaries. By providing this information to one set of learners and withholding it from another, it is possible to contrast learning that relies on transitional probabilities for word learning.

Given that natural language input contains multiple forms of information about word forms, it is possible that the nature of learning will change as learners accrue information from the input. The initial phase of learning may be characterized by different cognitive strategies than employed at even slightly later times during the learning process, when the learner has begun to accumulate experience with a language corpus. This was demonstrated in a recent natural-language study of Icelandic sentence forms (Plante, Patterson, Dailey, Almryde, & Fridriksson, 2014). This study demonstrated that patterns of activation can be highly dynamic with learning, even within the initial 10 minutes of exposure to a new language.

The present study uses functional magnetic resonance imaging (fMRI) to investigate the brain's response over time as adults are asked to identify individual words belonging to an unfamiliar language (Norwegian). Word segmentation has been examined previously using artificial languages (Cunillera et al., 2009; Karuza, Newport, Aslin, Starling, Tivarus, & Bavelier, 2013; McNealy, Mazziotta, and Dapretto, 2006; McNealy, Mazziotta, & Dapretto, 2010), which although informative, include the ecological limitations discussed above. Others have examined additional aspects of word learning based on preand post-learning scans, with the learning phase occurring outside the scanner (Callan, Callan, & Masaki, 2005; Golestani & Zatorre, 2004; Gullberg et al., 2010; Raboyeau, Marie, Balduyck, Gros, Démonet, & Cardebat, 2004; Raboyeau, Marcotte, Adrover-Roig, & Ansaldo, 2010; Newman-Norland, Frey, Petitto, & Grafton, 2006). However, these various methods can be more properly interpreted as examining the outcome of learning, rather than the learning process itself. We were interested instead in activation that occurs as the brain is encountering and engaged in learning from natural language input.

We know from the behavioral studies discussed above that adults show rapid word learning when stimuli contain high transitional probabilities for syllables belonging to individual words and low transitional probabilities for syllables in separate but adjacent words. Previous imaging studies of artificial languages have also capitalized on these stimulus properties to study word segmentation. However, no study to date has contrasted stimulus conditions in which strong cues to word boundaries were present or absent. This contrast is critical in order to clarify the role of brain regions identified previously in imaging studies of statistical learning. For example, some have suggested that tracking of transitional probabilities in the input is related to activation within the inferior frontal gyrus (Karuza et al., 2013). Others have made the case that temporal cortex is critical to statistical learning (McNealy et al., 2006; McNealy et al., 2010).

As discussed previously, natural languages contain properties in addition to transitional probabilities that assist in the identification of word units (e.g., prosodic patterns and phonotactic information). We know that adult learners are capable of simultaneous tracking of multiple sources of statistical information in the input (Romberg & Saffran, 2013).

Learners exposed to high predictability syllable transitions in a natural language context should be able to use transitional probabilities, as well as any of the other cues to word structure to aid identification of individual words. Indeed, information-rich input may promote the recruitment of a neural network that is quite extensive, and highly dynamic over time (cf. Plante et al., 2014). Furthermore, because use of statistical information has resulted in rapid learning of words in prior artificial language studies, learners provided with input in which transitional probabilities predict word boundaries should show better learning than those provided with low predictability input.

In contrast, learners provided with language input in which word units cannot be learned based on transitional probability may still be able to use other sources of information (e.g., phonotactics, prosodic patterns) to identify words. As noted previously, it is likely that these sources of information require more extensive experience with language input to be used effectively. Therefore, we predict that participants will be slower to identify individual words embedded in running speech in the absence of statistical cues, if this learning occurs at all. Critically, activation previously implicated in statistical learning in the context of artificial language learning should be less engaged when learners are provided with input that lacks strong statistical cues to word boundaries.

Because we anticipated that the active process of learning would be accompanied by a dynamic process of resource allocation over time (cf. Plante et al., 2014), we employed an analysis method well suited to understanding the pattern of coordination among brain regions as learning progresses. Independent Component Analysis (ICA) separates complex signals into simpler component signals that are statistically independent (see Beckman, 2012 or McKeown, Makeig, Brown, Jung, Kinderman, Bell, & Sejnowski, 1998 for overviews of the technique). The separate signals, or independent components (IC) are model-free estimates in that the analysis does not assume a shape or duration for the hemodynamic response. This is in contrast to use of a General Linear Model (GLM), in which all signal variation is tested against an ideal wave form intended to represent an *a priori* model of the shape and time course of the hemodynamic response. Because ICA separates the complex fMRI signal into component waveforms, this analysis can reveal correlated regions of activation, as well as separate time courses for activation that occurs within a single brain region (e.g., Karunanayaka, Schmithorst, Vannest, Szflarski, Plante, & Holland, 2010; Plante et al., 2014; Schmithorst, Holland, & Plante, 2006; Schmithorst, Holland, & Plante, 2007). Important to the design of this study, ICA is robust for detecting activation in response to auditory stimuli presented against background scanner noise (Langers & van Dijk, 2011). For these reasons, ICA tends to be more sensitive than GLM approaches for detecting significant signal changes in imaging data related to the cognitive components of the task (McKeown et al., 1998; Thornton, Rodionov, Laufs, et al., 2010).

To summarize, this study tests two major hypotheses. First, if listeners rely primarily on statistical learning mechanisms for segmenting words in natural languages, then words should be learned faster from input that contains statistical cues to word units than from input that lacks this information. Second, learners should differentially engage brain regions depending on whether they received input that contained strong statistical cues to word units. Only input that contains strong statistical cues to word units should robustly activate

ICA networks critical to statistical word learning. Other areas that may relate simply to language input in general should activate regardless of whether or not strong statistical cues are present.

2. Methods

2.1 Participants

Thirty-two adults participated in this study. Adults were randomly assigned to one of two learning conditions (High and Low Predictability). Sixteen adults (7 males, 9 females; mean age 24 years) were assigned to a 'High Predictability' condition. Eleven members of this group were right-handed, and five were left-handed. Sixteen adults (3 males, 13 females; mean age 25 years) were assigned to the 'Low Predictability' condition. Thirteen members of this group were right handed, and three were left-handed. We selected participants without regard to handedness because the majority of both right- and left-handed individuals have left lateralized language. This inclusion of both in the participant groups is consistent with representation of the general population. In particular, the language lateralization of these subjects was documented in a prior GLM analysis (Plante, Asbjørnsen, Almryde, & Patterson, in press). Although lateralization shifted leftwards with learning, all subjects showed leftward lateralization in classic language cortex from the start of the learning experiment.

All participants in both groups spoke English natively, and six considered themselves to be bilingual. These included 2 in the High Predictability condition (English-French, English-Korean) and 4 in the Low Predictability condition (English-Spanish, English-Japanese, English-Tagalog). However, none had prior exposure to Norwegian or spoke any closely related language (e.g., Dutch, Swedish). None reported neurological conditions (e.g., seizure disorder, traumatic brain injury), psychiatric conditions (e.g., depression, bipolar disorder), or developmental disorders that affect language processing (e.g., specific language impairment, learning disability, dyslexia). In addition, none had contra-indicators for MRI scanning. All provided informed consent according to procedures approved by the University of Arizona Institutional Review Board.

2.2 Stimuli

2.2.1 High Predictability Stimuli—The stimuli were those used in a previous behavioral experiment which demonstrated rapid learning of Norwegian words embedded in running speech (Kittleson et al., 2010). The focus of learning in that study, and in the present experiment, consisted of nine bi-syllabic Norwegian words (i.e., Target words), each of which was embedded into six different semantically and grammatically correct Norwegian sentences (one Target word per sentence). Sentences averaged six words in length (range 4 to 11 words) and were recorded by a male speaker of Norwegian in the Bergen dialect following the Bokmaal orthography. Sentences were spoken with the speaker's normal conversational speaking rate, using the prosody that would be expected for the structure and content of the sentences.

The Target words, embedded in the stimulus sentences, had an average forward transitional probability of .88 and an average two-way syllable dependency of .90. Additional statistical characteristics of the stimuli are available in Kittleson et al., 2010. The set of six sentences that each contained the same Target word were all presented within a single 18.2-second block within the fMRI scan. A brief period of silence separated each sentence. The length of this silent period varied according to how much time the nine individual sentences occupied in the 18.2-second block.

The position of the Target words within the sentences was varied. All Target words appeared in the initial position once. All but two words also appeared in the sentence final position. Words appearing in the initial position provided evidence to the learner that the initial syllable (when in sentence initial position) can sometimes occur independently of preceding syllables. Likewise, an appearance of a word in the final position provided evidence that the final syllable of Target words can sometimes occur independently of any syllables following the word. If learners use this evidence, it may reduce the task difficulty compared to learning solely embedded words. However, we chose to allow initial and final word positions in addition to embedded positions because words appear in all three positions in natural languages. For successful statistical learning, however, the learner would still need to recognize that the co-occurrence of two particular syllables constitutes a word unit.

2.2.2 Low Predictability Stimuli—The Low Predictability stimuli were derived from the same digital recordings used for the High Predictability condition. The Low Predictability sentences were edited by randomly selecting and clipping one syllable from the Target word for each High Predictability sentence. The clipped syllable was then inserted elsewhere, randomly, within the sentence stimuli. This served to break the statistical relation between the first and second syllable of the Target words. This manipulation lowered the transitional probability statistics for the syllables in the target words, while maintaining their overall frequency of occurrence in the input relative to the High Predictability condition. This rearrangement of syllables in the Low Predictability stimuli had the side effect of altering the prosodic contour of the sentence at the time points at which clipped syllables were inserted. However, given that the majority of each sentence was unaffected by these edits, these stimuli continued to strongly resemble a natural foreign language to a naïve listener. Importantly, since the majority of syllables in each sentence remained in their original order, the predominant (trochaic) Norwegian stress pattern was preserved. All other aspects of the Low Predictability condition were identical to the High Predictability condition.

2.2.3 Control Stimuli—Control stimuli were meant to reflect the general frequencies and dynamic changes in speech, without sounding speech-like. Accordingly, we created auditory stimuli by creating pure tones between the frequencies of 250 Hz and 2500 Hz of the same durations as the sentence stimuli (between 1800 and 3200ms). The frequency of these tones were modulated using Sound Forge 7.0 (Sony Pictures Digital Inc., 2003). A wide variety of frequency modulations were used across the set of tones so that each was acoustically distinct. Pairs of frequency-modulated tones of the same duration were then selected such that one tone was at an initial frequency of less than 1000 Hz and the other was at or above this frequency. These tones were then digitally merged to create a complex tone. Six of these

complex tones were selected at random for each control block. To mimic sentence blocks, tone stimuli were separated by brief periods of silence, which varied in length to make up the remainder of the 18.2-second control block period.

During the fMRI scans, sentences and control stimuli were presented as interleaved 18.2-second blocks. Each sentence and control block was preceded by a 2.6 second cue period in which the participant was instructed “Listen to these sentences” for the sentence block or “Just relax” for the control blocks. These cue stimuli were recorded by a female English speaker so that they were perceptually distinct from the Norwegian speaker. The fMRI stimuli began with the cue for the first control block, followed by control stimuli, the sentence cue, and sentence stimuli, with these blocks repeating for a total of nine blocks each. The total amount of time during which participants were exposed to the Norwegian stimuli was 2 minutes, 43.8 seconds during each fMRI scan.

2.2.4 Test Stimuli—A single set of test stimuli was used for both the High Predictability and Low Predictability groups. Test stimuli were taken from Kittleson et al. (2010) and consisted of 9 Target words and 9 Nonwords. It is important to note that in the Kittleson et al. study, individuals who had not been exposed to the Target words in sentences were unable to correctly identify Target vs. Nonwords at above chance rates in the Kittleson et al. study. This indicated that the Target words were no more “guessable” than the Nonwords.

Correct test items included the 9 two-syllable Target words, which were clipped from the sentences presented during scanning. Kittleson et al (2010) reported that the average frequency of occurrence of the correct Target words in Norwegian was 0.14 per 1000 words ($SD=0.33$). The forward transitional probability for the words in the sentence stimuli was 0.88 (range .46 to 1.0). The two-way syllable dependency (r_{ψ}) was .90 (range .47 – 1.0).

Target words were contrasted with Nonwords during the test phase. Nonwords were created by first clipping adjacent syllables from the recorded sentences. The syllable pairs consisted of one syllable from each Target word along with an immediately adjacent syllable. Four of these Nonwords contained the first syllable of the Target word and five contained the last syllable (refer to Kittleson et al., 2010 for these stimuli). Because of this composition, Nonwords contained one syllable that occurred as frequently in the sentence stimuli as the syllables of the correct Target words. However, the resulting bi-syllables were not real Norwegian words. As such, their frequency of occurrence as words in Norwegian was zero. As reported in Kittleson et al. (2010), the forward transitional probability for the Nonwords was .13 (range .08 to .20), and the two-way syllable dependency (r_{ψ}) was .18 (range .09 – .41). Thus, the statistical predictability of the syllable pairs for Nonwords in the training stimuli was much lower than for Target Words.

2.3 Procedures

2.3.2 Behavioral procedures—Prior to entering the scanner, each participant was informed of what to expect during the experimental portions of the session. They were told to attend to the stimuli during the scan and that they would be tested on what they learned after each scan. Prior to scanning, they were played an audio file that contained alternating blocks of the control tone stimuli and Norwegian sentences to practice the task they would

complete in the scanner. These practice stimuli did not contain any of the Target words or Nonwords that would be used during the test phase. During the fMRI scans, participants heard the same set of experimental stimuli during each of the three consecutive scans. Between scans, and while still in the scanner, participants listened to the test stimuli, administered in random order. Participants indicated, via button press, whether each test item was a real Norwegian word or not.

2.3.3 Imaging procedures—Structural and functional MRIs were collected on a General Electric 3 Tesla HD 12x Signa Excite system with an 8-channel head coil. Three fMRI scans were obtained using a spiral in/out sequence (Glover & Law, 2001) in the axial plane (TR 2600, TE 30, Flip angle 90, FOV 240, Matrix 64×64, 26 contiguous slices, voxel dimension 3.75×3.75×5mm). Each scan consisted of 173 volumes (including 9 prestimulus volumes and 2 post-stimulus volumes). Stimuli were arranged in blocks, beginning with the control stimuli. Sentence and control stimuli were heard in 9 blocks of 7 TRs (18.2 sec) each. A 2 TR (5.2 sec) cue period preceded each stimulus or control block. The total time for each scan was 7 minutes, 29.8 seconds, of which exposure to the Norwegian stimuli consisted of 2 minutes, 43.8 seconds.

A T1-weighted structural image in the same plane as the fMRI scan was collected (spin echo TR 500, TE minimum full, NEX 1, Phase FOV 1, FOV 240, Matrix 256×256, 26 slices, voxel dimensions 0.94×0.94×5mm). In addition, a high resolution SPGR image was obtained in the sagittal plane (TR=5.145ms, TE=1.988ms, inversion time=500ms, Flip angle 15, NEX 1, FOV 256mm, Matrix, 256×256 contiguous slices, voxel dimension of 1×1×1.4 mm).

2.3.4 Image analysis—Functional images were preprocessed using AFNI software (version AFNI_2011_12_21_1014, Cox, 2012). Images were evaluated to determine that the scanner had reached equilibrium by the stimulus onset. Images were registered to a base image selected for each subject to reflect minimal deviation from the average signal. The degree of movement was evaluated for each participant and averaged less than 0.4mm displacement. Two potential participants were excluded from the dataset due to excess movement leaving the 32 participants, as described above, in the final dataset. Slices in each image volume were time-shifted to the middle slice using Fourier interpolation. Each scan was despiked, and then normalized to a site-specific spiral template in standard space. Finally, data were smoothed with a 7 mm FWHM Gaussian kernel and the pre-stimulus scans were removed from the image set.

2.3.4.1 Independent Component Analysis: Independent component analysis was used to analyze the image data. We used GIFT (version 1.3i; Rachakonda, Egolf, Correa, & Calhoun, 2007) to perform the ICA analysis using the Extended Infomax algorithm, which has been shown to have a higher sensitivity to true signal change compared with other methods (Arya, Calhoun, Roys, Adali, Greenspan, & Gullapalli, 2003). The ICA analysis was run iteratively (10 times) so that stability of the ICA components could be evaluated using the *Icasso* option in GIFT. This is necessary because the stochastic nature of the ICA statistic can result in different estimations based on differences in the starting point of the analysis (Himberg, Hyvärinen, & Esposito, 2004). Separate analyses were conducted for the

High and Low Predictability Groups. For each analysis, subject data were concatenated (Schmithorst & Holland, 2004) and the scan number (Scan 1-3) was specified as a repeated measure. This promoted direct comparability of the resulting ICA components across individuals and across scans within each condition.

As with all multivariate statistical techniques applied to high dimensionality datasets, over-fitting of the data can result in identifying effects that are related to sample-specific and chance variation. In the case of ICA, over-fitting the data can reduce component stability. Conversely, under-fitting the data risks missing true effects, particularly if they are small in magnitude (Abou-Elseoud, Starck, Remes, Nikkinen, Trevonen, & Kiviniemi, 2010). Therefore, it is necessary to strike a balance between under- and over-fitting the fMRI data (Li, Adali, & Calhoun 2007). To address these issues, GIFT can be used to estimate the optimum number of components for a given dataset (See Li et al., [2007] for a complete description of the algorithm and its effectiveness). We used this estimation procedure and allowed for 40 components to be generated during the analysis. Although it is computationally time consuming, we used the full set of subjects to estimate the optimal number of components, which yields the best possible estimate. The 40 components estimated was in excess of the number of meaningful, task related components expected for each group. However, this number also allowed for components related to physiologic and other noise to be estimated by the ICA analysis. The resulting 40 ICA waveforms were then regressed against the fMRI signal to produce 40 image maps. These were thresholded to a brain-wide family-wise error (FWE) corrected value of $p < .05$ for display.

Of the forty ICA components, we selected those for which signal was positively correlated with the sentence blocks ($p = .05$ or better) for at least one of the three scans. This criterion was applied so that brain regions represented by each component that might change dynamically with learning were considered for all three consecutive scans, even if the entire component did not meet the significance criterion for all phases of learning. This significance criterion reduced the number of components for consideration from 40 to 11 for the High Predictability group and 40 to 4 components for the Low Predictability group. These components were visually inspected to determine which maps showed characteristics of artifacts (e.g., movement artifacts, susceptibility artifacts, fluid movement) as ICA estimates signal variation regardless of its source. Movement in particular, whether muscular or physiological, can be identified as a signal source in ICA even if data have been motion-corrected during preprocessing (McKeown, Hansen, & Sejnowski, 2003). Evaluation of images for probable artifacts further reduced the number of viable components to 5 components for the High Predictability group and 2 components for the Low Predictability group. These components were evaluated for the stability of the ICA estimate. Each of these ICA components was highly stable (Iq values of .97 or greater).

Once ICA components of interest were identified, we were further interested in how subregions within the ICA component may have shifted over the course of learning. To extract these subregions, we thresholded the t-statistic images at a FWE corrected value of $p < .05$ using AFNI's 3dClustSim program. We then eroded the images at a threshold of 5% of the maximum t-value. This removed diffuse regions of low-level activation and helped to separate adjacent regions that contained separate peak activations. Visual inspection of the

resulting data indicated that this procedure separated a majority of peaks that appeared to represent separate anatomical regions. However, there were additional regions that contained two or more peaks. Typically, these were activation peaks that occurred close together in the three-dimensional space so that the intensity drop off between the peaks did not fall below the 5% erosion criterion. We separated these adjacent peaks into distinct regions only when the peaks clearly belonged to separate anatomical regions (e.g., the inferior occipital lobe and superior aspect of the cerebellum; inferior frontal gyrus and anterior temporal lobe; inferior frontal gyrus and middle frontal gyrus). We used the probabilistic anatomical atlases available in FSLView version 3.1 (FSL, Jenkinson, Beckman, Behrens, Woolrich, Smith, 2012), primarily the Harvard-Oxford Cortical and Subcortical Structural Atlases and the Probabilistic Cerebellar Atlas to guide this determination. Variations in activation strength that were more likely to lie within a single anatomical region (e.g., large activations centered on the superior temporal gyrus) were not separated further. We then considered the peak activation values within each of these regions.

2.3.4.2 General Linear Model Analysis: A general linear model analysis of the preprocessed data was also conducted using AFNI. Data were initially smoothed with a 7mm FWHM kernel. Signal intensities were fitted to the AFNI default model for BOLD activation on a subject-by-subject basis. Activation for both the High and Low Predictability groups was averaged and thresholded ($p < .05$ FWE) to correct for multiple comparisons over the full brain volume.

3. Results

3.1 Behavioral Results

Figure 1 displays the group behavioral results as well as an indication of variability within and across individuals in both the High (Fig. 1A) and Low Predictability groups (Fig. 1B). Learning was operationally defined as significantly greater acceptance of Target words compared with Nonwords. We tested the relative acceptance rate for Target words and Nonwords across time using a mixed ANOVA (Statistica, v8.0 Statsoft, 2008) in which High Predictability and Low Predictability conditions were the between-group factor and item (acceptance rate for Target and Nonwords) and time (Test 1-3) were within-subject conditions. This type of analysis is typical of the artificial language learning literature because it controls for any systematic bias in the base rate of acceptance. There were main effects for Item ($F(1,30)=39.72$, $p=.000001$) and Time ($F(2,60)=10.53$, $p=.00012$) and a significant Item x Time interaction ($F(2,60)=5.72$, $p=.00534$). There were no main effects or interactions involving group. However, as Figure 1 indicates, there was a trend towards group differences in acceptance of Target vs. Nonwords, particularly at Test 1 and 2.

We hypothesized that High Predictability input would result in faster learning than Low Predictability input. Least squares analyses indicated significant learning in the High Predictability group at all three test times ($F(1,30)=4.42$, $p=.044$; $F(1,30)=12.82$, $p=.001$; $F(1,30)=23.64$, $p=.00003$ for Test 1, 2, and 3 respectively). For them, above chance performance was present after the initial scan (2 minutes, 43.8 seconds of exposure) and the

increase in accuracy was a fairly linear function of time. The Low Predictability group accepted more Target words than Nonwords only at Test 3 ($F(31.10, p=.000005)$), after 8 minutes, 11.4 seconds of exposure. Item responses were not different at Tests 1 and 2 ($F(1,30)=.55, p=.46$; $F(1,30)=2.45, p=.13$ respectively). Performance by the Low Predictability group at Test 3 suggests that this group may have gleaned information other than syllable co-occurrence (e.g., prosodic patterns, phonotactic information) that assisted them in identifying Target words, but not until after the third scan. These data also show that group differences emerged, despite noteworthy inter- and intra-subject variability in the data.

Because some of our participants were bilingual, we considered the idea that those with fluency in multiple languages might be more proficient learners than their monolingual counterparts. The numbers of bilingual participants within each condition were too small to permit statistical testing of subgroup differences. However, we examined the scores of the six bilingual participants against the range of scores for their participant group. These participants completed a total of 18 post-scan tests during the experiment. Overall, the majority of these 18 test scores (72% overall) were within 1SD of their monolingual counterparts. Only one bilingual (in the Low Predictability group) scored consistently at the high end (>1 SD above the monolingual mean) on all three post-scan tests. However, the remaining three bilinguals in this condition were consistently within 1SD of the monolingual mean. Therefore, being bilingual did not appear to confer any overall learning advantage in this study.

3.2 Imaging Results

3.2.1 Overview of Independent Components—The independent component analysis yields collections of regions for which signal changes occur over the same time course. Therefore, these regions can be interpreted as showing functional connectivity. The analysis procedures yielded 5 task-related components for the High Predictability group and 2 task-related components for the Low Predictability group. For these ICs, b values for the entire IC, averaged across all its voxels, were statistically significant ($p<.05$) during at least one of the three scans. Figure 2A and B provide a graphic representation of the change in b values for each ICA component over time. As this figure demonstrates, the different IC signals showed different levels of activation, and those showed different patterns of change across time.

The brain-wide activation contained within each IC was thresholded at a FWE corrected p value of .05. The individual regions of activation this produced are shown in Figure 3A (High Predictability condition) and Figure 3B (Low Predictability condition) and are color-coded for their IC. This figure provides an overview of the spatial change in activation for each IC over time.

To describe ICs in terms of regional activation, the IC-wide activations were eroded by 5% as described in the Methods. This served to separate areas of peak activation within each IC. The location (in MNI coordinates supplied by FSL), extent of activation (number of voxels), and strength of the peak activation (t statistic) in each region is provided in Supplemental Table 1. It is clear from the Figure 3 and Supplemental Table 1 that the High Predictability

condition not only engaged more ICs, but also more regions within each IC than did the Low Predictability condition.

As Supplemental Table 1 indicates, there were few areas of overlap among the different ICs for the High Predictability group. The major regions of overlap between ICs are displayed in Supplemental Figure 1. Regional overlap provides a potential mechanism by which activation associated with different ICs might be integrated into the overall network. As Supplemental Figure 1 indicates, there were several areas of overlap along the superior temporal sulcus, the posterior temporal lobe, and within the lateral temporal-parietal-occipital junction for the High Predictability Group. There was no regional overlap for ICs obtained for the Low Predictability group (see Supplemental Table 1 and Figure 3B).

The only region of activation in common for both the High and Low predictability groups was in and around the superior temporal sulcus. A direct comparison of ICA-derived activation in this region for the High and Low Predictability group was not possible for several reasons. First, because separate ICAs were conducted for each group, there is no guarantee that waveforms generated by each ICA reflected equivalent signal in terms of time course, even when they came from an equivalent brain regions. Furthermore, the ICA data indicated qualitative differences in activation between groups (see Figure 3A &B). This region was active solely in IC1 for the Low Predictability group. In contrast, multiple ICs included activation in this region for the High Predictability group. This means that this region activated only on one time course (i.e., one hemodynamic response) for the Low Predictability group and had multiple, independent time courses (i.e., multiple hemodynamic responses) for the High Predictability group. Thus, there was not a single signal that could serve as a common basis for comparison between the two groups.

For a direct comparison of activation in the temporal lobe region, it was necessary to obtain a single waveform that could be compared directly between groups. To accomplish this, we used temporal lobe activation data derived from a separate GLM analysis. Because the GLM analysis assumes a common, a priori waveform, GLM data can be compared directly between groups. The thresholded brain activation revealed by the GLM analysis was used to develop a data-driven temporal region mask. This mask was based on the combined activation seen for the two groups and three scans so that activation data was extracted for a common anatomical region for both groups and across time.

The brain images displayed in Figure 4 displays the brain-wide results of the GLM analysis, for each of the three scans. The accompanying graphs plot the average t-statistic within the right and left temporal regions for each scan and each group. A mixed ANOVA was performed on this data, with learning condition (High vs. Low Predictability) as a between-group factor and hemisphere and scan as within group factors, and the b-values as the dependent measure. The ANOVA revealed a significant main effect for scan ($F(2,60)=9.25$, $p=.003$), with activation generally becoming stronger over time. There was also a significant scan x hemisphere effect ($F(2,60)=4.03$, $p=.023$) and scan x hemisphere x learning condition effect ($F(2,60)=3.93$, $p=.025$). The trend for the three-way interaction was for the High predictability group to show greater change within the left hemisphere over time. However,

post-hoc analysis failed to find any pair-wise differences when corrected for multiple comparisons.

3.2.2 Brain-behavior correlations—We correlated behavioral performance with regional activation obtained from the ICA data. We used the back reconstruction algorithm from GIFT to produce waveforms for each subject corresponding to the different ICs. We then developed IC-specific masks, based on the spatial distributions of significant activation for each IC (shown in Figure 3). These were used to extract IC-related activation on an participant-by-participant basis. These IC-based activations were further divided into subregions that occurred within each lobe, the insula, cerebellum, and subcortical areas in the right and left hemisphere for each IC. Note that not all of these subregions were present for each IC. We calculated separate correlations for correct accept responses and for correct reject responses because these two types of responses require different cognitive skills (i.e., confirmation vs. error detection).

For the High Predictability group, significant positive correlations ($p < .05$) were found between correct accept responses post scan and activation in the left dorsolateral prefrontal ($r = .57$), and right temporal lobe ($r = .56$) regions of IC1 during scan 1. In these cases, better detection of correct test items was associated with stronger activation in these cases. Significant negative correlations during Scan 1 were also found for this group for correct reject responses. These included the left ($r = -.57$) and right ($r = -.57$) frontal lobe regions, the left ($r = -.55$) and right ($r = -.59$) temporal lobe regions, and the left parietal including the angular and supramarginal regions ($r = -.57$) of IC3. In these cases, better ability to reject incorrect test items post-scan was associated with lower activation in these regions during the preceding scan.

No significant correlations were found for Scans 2 or 3 and no correlations were found for the Low Predictability group for any of the three scans. The lack of behavioral correlations for the High Predictability group in the later scans, or for the Low Predictability group, who did exhibit learning by Scan 3 should be interpreted with caution. The small range of possible scores (0-9 each for correct accepts and correct rejects) likely reduced the probability of detecting correlations. Indeed, the test results following Scan 1 showed the widest range of behavioral scores (2 to 9 for correct accepts and 2 to 7 for correct rejects) of any of the three test sessions.

4. Discussion

Consistent with the idea that learning based on the presence of statistical relations in the input results in rapid, unguided learning (Gómez, 2006; Newport & Aslin, 2004; Saffran, 2003), participants who received input with strong statistical cues to word units learned the Target words after relatively little exposure to Norwegian sentences. Somewhat surprisingly, participants who were given input that lacked strong statistical cues also learned to differentiate between real and pseudo-words, although the time needed was three times as long. This difference in time required to learn is consistent with the idea that participants in the Low Predictability group were using information other than statistical regularities to make their decisions about whether test items were real or pseudo-words.

Activation to high and low predictability input was plainly different. Surprisingly, this was true even during the first scan when both groups had minimal exposure to the Norwegian language stimuli. From the first scan, those exposed to stimuli constructed to facilitate rapid learning of the Target words promoted the engagement of a network comprised of five functionally connected sub-networks, each of which operated on a distinct time course (IC). In contrast, low predictability input evoked a more limited response with only one network active in Scan 1 and a second network added in Scans 2 and 3.

4.1 Statistical input as an attentional magnet

The striking difference in the IC networks for the two groups occurred despite the fact that all participants heard nearly identical stimuli. These results clearly demonstrate that word learning, when syllable-level statistics are available, is neurobiologically distinct from when this type of information is not available to learners. The difference is consistent with the notion that statistical regularities draw attention in ways that unstructured input does not. Independent evidence for this idea comes from a series of studies (Zhao & Turk-Browne, 2013) in which visual arrays were shown to learners. Although only a subset of the visual input contained statistical dependencies and participants were engaged in an unrelated cover task, participants nonetheless showed learning of the statistical dependencies. In the present study, many of the regions active under the High Predictability condition are ones that have been previously associated with the engagement of attentional resources for language stimuli.

Auditory attention is necessary for the recognition of word forms. The anterior cingulate cortex, implicated for general attentional functions, activated consistently in the left hemisphere for all scans in the High Predictability condition but only during Scan 1 in the Low Predictability condition. The lack of continued activation under the Low Predictability condition may well reflect difficulty in developing or maintaining a particular attentional strategy over time, given the weak statistical relations among syllables in this condition. For the Low Predictability group, syllables contained in the Target words lacked the predictive power to draw the listener's attention. In contrast, the High Predictability group could learn to anticipate the second syllable of each Target word once the first was heard. This likely affected the degree to which attention was attracted by the stimuli for two groups.

A focus of attention on phonological sequences would be expected to involve regions more particularly associated with auditory attention. Regions in the temporo-parietal-occipital junction are likely to serve this function. The High Predictability condition only produced large activations in the left posterior middle temporal and angular gyri, extending into the lateral occipital and supramarginal gyri. The left angular and inferior supramarginal gyri are known to activate for acoustic contrasts that signal phonological distinctions (Turkeltaub & Coslett, 2010). These regions also show greater activation for words compared to pseudo-words (Price, 2010). This suggests that this general region is less tuned to specific auditory stimuli (e.g., sounds vs. words) than to placing the auditory stimuli within the focus of attention (Chein, Ravizza, & Fiez, 2003; Christensen, Almryde, Fidler, Lockwood, Antonucci, & Plante, 2012). The supramarginal gyrus is more sensitive to whether words were previously within the focus of the listener's attention and the angular gyrus responds

more to words that had not been heard previously during the experiment (i.e., not previously attended at all) (Christensen et al., 2012). In the present study, activation within the supramarginal and angular regions was negatively correlated with correct reject responses. Consistent with the Christensen et al. (2012), this can be explained by the likelihood that Target and Nonwords embedded in the input were drawn into the learner's attentional focus. The low statistical predictability of Nonwords should not have drawn these words into the focus of attention, resulting in lower activation posterior cortex. The lower the activation, the less likely that the syllable co-occurrences comprising Nonwords would be identified as real words at test.

Other roles have been attributed to the temporal-parietal region that are potentially relevant to word learning, but are a more tenuous fit for the present study. For example, others have pointed to regions within the temporal-parietal junction as the locus of phonological rehearsal (cf. Baddeley & Hitch, 1974; see Chien et al., 2003 for a review of imaging data). However, performance on our task is unlikely to have capitalized on a phonological loop-type mechanism. The number and speed with which individual syllables were presented within the sentence stimuli would have exceeded most listeners' capacity for active rehearsal. In addition, the angular and supramarginal gyri were only active during the High Predictability condition, indicating that only the repeated co-occurrence of syllables comprising Target words differentially engaged these regions. This would be the case if the highly predictable nature of the Target words served to draw these particular sequences into the focus of attention during learning.

The inferior parietal region, along with the temporal lobe, is consistently active in semantic processing tasks (Binder & Desai, 2011; Price, 2010) and the inferior parietal regions in particular become better integrated into a larger word-learning network for learners who are able to successfully associate phonological strings with meaning (Yang, Gates, Molenaar, & Li, 2015). However, our task did not provide learners with any semantic information for the words contained in the input. Therefore, the extent of lexicalization in our study would necessarily be limited to recognizing particular phonological strings as a word unit. However, it is unlikely that the inferior parietal activation contributed to this aspect of the lexicalization process. This region was not active for the Low Predictability group even after this group demonstrated above-chance identification of Target words. This argues against a more general role of this area in early-stage lexicalization in which phonological strings become word forms.

It is worth noting that the supramarginal gyrus within the inferior parietal region did not have a highly consistent time course of activation across scans, even though it was consistently active. In addition, a large number of other cortical and subcortical regions activated only transiently across the three scans of the experiment. The transient appearance of some regions and the shifts in functional connectivity for others has been reported previously when learning is imaged over time (e.g., Bassett, Wymbs, Porter, Mucha, Carlson, & Grafton, 2011; Plante et al., 2014). These types of results suggest a highly dynamic process of resource allocation as the brain 'learns how to learn'. The fluctuations in the overall pattern of activation may reflect implicit processes as the brain seeks to determine which aspects of the input are informative. This may lead to the temporary

recruitment of regions, which are later abandoned when they prove unproductive relative to the task at hand. In addition, learners may adopt strategies, which may be implicit or explicit, that are applied at different stages of learning (e.g., track transitional probabilities, recognize familiar phones, attend to stress patterns). This would also likely contribute to the dynamic pattern of activation seen within the experiment.

4.2 Comparison to artificial language studies

The most highly consistent region of activation reported for word segmentation in artificial languages has been the superior temporal gyrus (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2006; McNealy et al., 2010). In the present study, this region activated to both high and low predictability input. However, there were several group differences related to activation of this region. For the High Predictability group only, the left, but not right temporal region correlated with correct accept responses during Scan 1. This does not necessarily reflect the lack of a role for the left temporal region, but may instead reflect the degree to which successful learners in the High Predictability group also recruited the right hemisphere homolog in order to exhibit early learning. The Low Predictability group did not evidence learning during Scan 1. In addition, the extent and strength of temporal activation was fairly comparable in this region during the initial scan, but activation diverged with time under the two input conditions. The contrasting pattern of temporal lobe activation in the Low and High Predictability conditions suggests that this particular region contributes to the tracking of statistical information (Hickok, 2012), as well as to more basic aspects of language processing.

In prior artificial language studies of word segmentation, frontal activation has rarely been reported. However, one study (Karuza et al., 2013) has suggested that the left inferior frontal gyrus contributes specifically to tracking statistical information during word learning. Although this region was not significantly active in their whole-brain GLM analysis, they found a significant correlation for regional activation with a behavioral metric that reflected differences in learner performance across time points. We found a positive correlation between correct accept performance and frontal lobe activation for the High Predictability group only. This occurred only for IC1 (for Scan 1), in which significant dorsolateral prefrontal rather than the inferior frontal activation was found. This correlation suggests a contribution to control of cognitive resources rather than to word segmentation.

Our GLM analysis also detected the dorsolateral prefrontal activation, but inferior frontal gyrus activation was not statistically significant. However, the more sensitive ICA technique revealed inferior frontal activation, and only in the High Predictability condition. However, the inferior frontal activation was not consistently present across time. Pars opercularis in the left hemisphere was active (IC3 and IC4), but only during Scans 1 and 2. Pars orbitalis (IC4) was active during Scan 3 only. It might be that pars opercularis was only engaged during the first two scans because learners in the High Predictability group no longer needed to rely as heavily on distributional information by Scan 3. Although possible, this explanation is unsatisfying in light of the behavioral data. The High Predictability group was well below ceiling performance at the end of the experiment, indicating that participants were still in the process of learning to segment words from the speech stream. The fact that

the left inferior frontal gyrus was not consistently active in this behavioral context reduces the likelihood that its primary role is that of tracking statistical information in the input.

Activation in subcortical regions is inconsistently reported for artificial word segmentation paradigms. For example, GLM analyses have detected basal ganglia activation in studies by McNealy and colleagues (McNealy et al. 2006; 2010), but not by others who used highly similar stimuli (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2010). Karuza et al (2013) reported correlations between their change metric and basal ganglia activation, but only when the basal ganglia was masked specifically for the analysis. In the present study, left putamen, right caudate, and left globus pallidus activation were detected at various times during the learning period. Of these, only the caudate activated for both the High and Low Predictability groups. The timing of this activation within the experiment was consistent with the onset of behaviorally measurable learning for each group. For the High Predictability group, right caudate activation occurred in Scan 1, and this group demonstrated learning during the immediately following test phase. For the Low Predictability group, right caudate activation occurred in Scan 2, just preceding evidence of learning (obtained after Scan 3).

For both groups, the timing of this activation is generally consistent with the idea that caudate activity may contribute to early stages of learning. Our findings have a parallel in a visual statistical learning study (Turk-Browne, Scholl, Chun, & Johnson, 2009). That study reported right caudate activation after 48 seconds of exposure to structured visual input and was not detected at later time points in the experiment. The onset of behaviorally-measurable learning for this study was not known. However, it was not likely to precede 48 seconds, because behavioral performance measured at the end of the experiment indicated that participants were still in the process of acquiring the statistical patterns.

4.3 Conclusions

In summary, the view of learning presented by our natural language task includes both points of convergence and divergence with previous studies that employed artificial languages to study word segmentation from running speech (Cunillera et al., 2009; Karuza, et al., 2013; McNealy et al., 2006 2010). In all studies, including the present study, activation of the superior temporal gyrus has been reported. By using an analysis technique that is sensitive to activations corresponding to independent time courses, we also show that this area, along with additional areas in the temporo-parietal-occipital junction and frontal regions activate very differently depending on whether or not the stimuli contain the types of predictable relations that facilitate statistical learning.

It is important to note, however, that our natural language task is not directly comparable to the previous studies of statistical learning in all respects. The Norwegian words used as input in the present study occurred at a much lower frequencies than is typical of the analogous artificial language studies. Furthermore, our American learners encountered Norwegian phonemes that were foreign to them. These factors may have led our learners to rely more on additional posterior auditory and frontal regions, compared with previous artificial learning studies. These regions may be more useful for learning strategies required

for the characteristics of an unfamiliar natural language than those required to process the highly constrained and phonologically familiar stimuli of artificial languages.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This work was supported by NIH Grant R01 DC011276.

6. References

- Abou-Elseoud A, Starck T, Remes J, Nikkinen J, Tervonen O, Kiviniemi V. The effect of model order selection in group PICA. *Human Brain Mapping*. 2010; 31:1207–1216. [PubMed: 20063361]
- Adriaans F, Kager R. Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language*. 2010; 62:311–331.
- Aguilar JM, Plante E. Learning of grammar-like visual sequences by adults with and without language-learning disabilities. *Journal Of Speech, Language, And Hearing Research*. 2014; 57(4): 1394–1404.
- 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. p. 1-1.
- Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*. 1998; 9:310–324.
- Baddeley, AD.; Hitch, G. Working memory. In: Bower, GH., editor. *The psychology of learning and motivation*. Vol. 8. Academic Press; New York, NY: 1974. p. 47-89.
- Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, Grafton ST. Dynamic reconfiguration of human brain networks during learning. *PNAS Proceedings Of The National Academy Of Sciences Of The United States Of America*. 2011; 108:7641–7646.
- Binder JR, Desai RH. The neurobiology of semantic memory. *Trends in Cognitive Sciences*. 2011; 15:527–536. [PubMed: 22001867]
- Beckman CF. Modelling with independent components. *NeuroImage*. 2012; 62:891–901. [PubMed: 22369997]
- Bortfeld H, Morgan J, Golinkoff R, Rathbun K. Mommy and me: Familiar names help launch babies into speech0stream segmentation. *Psychological Science*. 2005; 16:298–304. [PubMed: 15828977]
- Brady TF, Oliva A. Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science*. 2008; 19(7):678–685. [PubMed: 18727783]
- Brent M, Siskind J. The role of exposure to isolated words in early vocabulary development. *Cognition*. 2001; 81:33–44.
- Buckner RL. Functional-anatomic correlates of control processes in memory. *The Journal of Neuroscience*. 2003; 23:3999–4004. [PubMed: 12764084]
- Callan AM, Callan DE, Masaki S. When meaningless symbols become letters: Neural activity change in learning new phonograms. *NeuroImage*. 2005; 28:553–562. [PubMed: 16055350]
- Chein JM, Ravizza SM, Fiez JA. Using neuroimaging to evaluate models of working memory and their implications for language processing. *Journal of Neurolinguistics*. 2003; 16:315–339.
- Christensen TA, Almryde K, Fidler LJ, Lockwood JL, Antonucci SM, Plante E. Modulating the focus of attention for spoken words at encoding affects fronto-parietal activation for incidental verbal memory. *International Journal of Biomedical Imaging*. 2012; 2012:579–586.
- Conway CM, Christiansen MH. Modality-Constrained Statistical Learning of Tactile, Visual, and Auditory Sequences. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*. 2005; 31(1):24–39.

- Cox RW. AFNI: What a long strange trip it's been. *NeuroImage*. 2012; 62:743–747. [PubMed: 21889996]
- Cunillera T, Càmarà E, Laine M, Rodríguez-Fornells A. Words as anchors: Known words facilitated statistical learning. *Experimental Psychology*. 2010; 57:134–141. [PubMed: 20178930]
- Cunillera T, Càmarà E, Torro JM, Marco-Pallares J, Sebastián-Galles, Ortiz H, Pujol J, Rodríguez-Fornells A. Time course and functional neuroanatomy of speech segmentation in adult. *NeuroImage*. 2009; 48:541–553. [PubMed: 19580874]
- Cutler A, Norris D. The role of strong syllables in segmentation for lexical access. *Journal Of Experimental Psychology: Human Perception And Performance*. 1988; 14(1):113–121.
- 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]
- Evans JL, Saffran JR, Robe-Torres K. Statistical learning in children with specific language impairment. *Journal of Speech, Language, & Hearing Research*. 2009; 52:321–335.
- Fernald A, Morikawa H. Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*. 1993; 64:637–656. [PubMed: 8339686]
- Fiser J, Aslin RN. Statistical learning of higher-order temporal structure from visual shape sequences. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*. 2002; 28(3):458–467.
- Fiser J, Aslin RN. Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*. 2001; 12(6):499–504. [PubMed: 11760138]
- 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]
- Golestani N, Zatorre RJ. Learning new sounds of speech: reallocation of neural substrates. *NeuroImage*. 2004; 21:494–506. [PubMed: 14980552]
- Gómez, RL. Dynamically guided learning. In: Munakata, Y.; Johnson, M., editors. *Attention & Performance XXI: Processes of change in brain and cognitive development*. Oxford University Press; Oxford, UK: 2006. p. 87-110.
- Graf Estes K, Evans JL, Alibali M, Saffran JR. Can infants map meaning to newly segmented words? *Psychological Science*. 2007; 18:254–260. [PubMed: 17444923]
- Gullberg M, Roberts L, Dimroth C, Veroude K, Indefrey P. Adult language learning after minimal exposure to an unknown natural language. *Language Learning*. 2010; 60(Suppl. 2):5–24.
- Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004; 82:67–99. [PubMed: 15037127]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci*. 2007; 8:393–402. [PubMed: 17431404]
- 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]
- Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012; 62:782–790. [PubMed: 21979382]
- Johnson E, Tyler M. Testing the limits of statistical learning for word segmentation. *Developmental Science*. 2010; 12:339–345. [PubMed: 20136930]
- Jusczyk PW, Cutler A, Redanz NJ. Infants' preference for the predominant stress patterns of English words. *Child Development*. 1993; 64(3):675–687. doi:10.2307/1131210. [PubMed: 8339688]
- Kittleson M, Aguilar JM, Tokerud GL, Plante E, Asbjørnsen AE. Implicit Language Learning: Adult's ability to segment words in Norwegian. *Bilingualism, Language, & Cognition*. 2010; 13:513–523.
- 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 EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D. The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain & Language*. 2013; 127:46–54. [PubMed: 23312790]
- Langers DRM, van Dijk P. Robustness of intrinsic connectivity networks in the human brain to the presence of scanner noise. *NeuroImage*. 2011; 55:1617–1632. [PubMed: 21255663]
- Li Y-O, Adali T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*. 2007; 28(11):1251–1266. [PubMed: 17274023]
- 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]
- McKeown MJ, Makeig S, Brown GG, Jung T, Kinderman SS, Bell AJ, Sejnowski TJ. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*. 1998; 6:160–188. [PubMed: 9673671]
- McNealy K, Mazziotta JC, Dapretto M. Cracking the language code: Neural mechanisms underlying speech parsing. *The Journal of Neuroscience*. 2006; 26:7629–7639. [PubMed: 16855090]
- McNealy K, Mazziotta JC, Dapretto M. The neural basis of speech parsing in children and adults. *Developmental Science*. 2010; 13:385–406. [PubMed: 20136936]
- Newman-Norlund RD, Frey, Petitto L-A, Grafton ST. Anatomic substrates of visual and auditory miniature second-language learning. *Journal of Cognitive Neuroscience*. 2006; 18:1984–1997. [PubMed: 17129186]
- Newport EL, Aslin RN. Learning at a distance: Statistical learning of non-adjacent dependencies. *Cognitive Psychology*. 2004; 48:127–162. [PubMed: 14732409]
- Pelucchi B, Hay JF, Saffran JR. Statistical learning in a natural language by 8-month-old infants. *Child Development*. 2009; 80:674–685. [PubMed: 19489896]
- Plante E, Asbjørnsen AE, Almryde K, Patterson DK. Language lateralization shifts with learning by adults. *Laterality: Asymmetries of the Body, Brain, & Cognition*. (in press).
- Plante E, Patterson D, Dailey NS, Almryde K, Fridriksson J. Dynamic Changes in Network Activations Characterize Early Learning of a Natural Language. *Neuropsychologia*. 2014
- Price CJ. 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]
- Raboyeau G, Marie N, Balduyck S, Gros H, Démonet J-F, Cardebat D. Lexical learning of the English language: A PET study in healthy French subjects. *NeuroImage*. 2004; 22:1808–1818. [PubMed: 15275937]
- Rachakonda, S.; Egoft, 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)
- Raboyeau G, Marcotte K, Adrover-Roig D, Ansaldo AI. Brain activation and lexical learning: The impact of learning phase and word type. *NeuroImage*. 2010; 49:2850–2861. [PubMed: 19837173]
- 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. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*. 2003; 12:110–114.
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science*. 1996; 274:1926–1928. [PubMed: 8943209]
- Saffran JR, Newport EL, Aslin RN. Word segmentation: The role of distributional cues. *Journal of Memory and Language*. 1996; 35:606–621.
- Saffran JR, Newport EL, Aslin RN, Tunick RA, Barrueco S. Incidental language learning: Listening (and Learning) out of the corner of your ear. *Psychological Science*. 1997; 8:102–105.
- 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. 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]
- Statsoft. Inc.. Statistica (v. 8.0). Author; Tulsa, OK: 2008.
- Sony Pictures Digital, Inc.. Sound Forge 7.0. Author; Madison, WI: 2003.
- Thiessen ED. Effects of visual information on adults' and infants' auditory statistical learning. *Cognitive Science*. 2010; 34:1093–1106. [PubMed: 21564244]
- Thiessen ED, Saffran JR. When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9- month-old infants. *Developmental Psychology*. 2003; 39:706–716. [PubMed: 12859124]
- Thorton RC, Rodionov R, Laufs H, Vulliemoz S, Vaudano A, Carmichael D, Cannadathu S, Guye M, McEvoy A, Lhatoo S, Bartolomei F, Chauvel P, Diehl B, De Martino F, Elwes RCD, Walker MC, Duncan JS, Lemieux L. Imaging haemodynamic changes related to seizures: Comparison of EEG-based general linear model, independent component analysis of fMRI and intracranial EEG. *NeuroImage*. 2010; 53:196–205. [PubMed: 20570736]
- Turk-Browne NB, Jungé JA, Scholl BJ. The Automaticity of Visual Statistical Learning. *Journal Of Experimental Psychology: General*. 2005; 134(4):552–564. [PubMed: 16316291]
- Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM. Implicit perceptual anticipation triggered by statistical learning. *The Journal Of Neuroscience*. 2010; 30(33):11177–11187. [PubMed: 20720125]
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK. Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*. 2009; 21:1934–1945. [PubMed: 18823241]
- Turkeltaub PE, Coslett HB. Localization of sublexical speech perception components. *Brain & Language*. 2010; 114:1–15.
- Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*. 2006; 30:1414–1432. [PubMed: 16413796]
- Yang C. Universal grammar, statistics or both? *Trends in Cognitive Sciences*. 2004; 8:451–456. [PubMed: 15450509]
- Yang J, Gates KM, Molenaar P, Li P. Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics*. 2015; 33:29–49.
- Zhao J, Al-Aidroos N, Turk-Browne NB. Attention is spontaneously biased toward regularities. *Psychological Science*. 2013; 24(5):667–677. doi:10.1177/0956797612460407. [PubMed: 23558552]

Highlights

- The study provides a natural language application of artificial language learning paradigm.
- Learning is visualized as it occurs rather than measured after a period of training.
- Learning associated with the presence of high transitional probabilities in the input is neurobiologically different than when the input lacks this fundamental characteristic.
- ICA analysis reveals the dynamics of early learning period.

Figure 1A & B.

Behavioral performance for the A) High Predictability and B) Low Predictability groups.

Behavioral results of a post-scan test indicated learning had occurred after all three scans for the High Predictability Group and after Scan 3 for the Low Predictability group.

Figure 2A and B.
Changes in the strength of activation for IC components in the A) High Predictability Condition and B) Low Predictability Condition.

Figure 3A & B.

Areas of significant activation ($p < .05$ FWE correction) for A) the five independent components in the High Variability condition and B) the two independent components in the Low Variability condition.

Figure 4. Comparison of activation for the High and Low Predictability groups using a GLM analysis. Graphs display the beta values in the temporal lobe regions over time.

Table 1

Regions of activation for High and Low Predictability conditions.

High Predictability		Scan 1					Scan 2					Scan 3				
IC	ROI	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z
1	L Cerebellum-vermis						38	3.66	-4	-68	-18					
	Cingulate-posterior											51	3.49	2	-44	34
	PostCG-dorsolateral											156	3.96	-52	-20	54
	PreCG-dorsolateral						711	4.57	-28	-2	44	689	3.74	-34	-2	50
	SFG-dorsal/ paracingulate						329	4.16	-16	6	42					
	STG	1817	3.65	-58	-42	8	4547	4.65	-58	6	-6	4718	4.69	-56	-38	18
	Thalamus						225	4.33	-8	-30	8					
	R Cerebellum-VI						352	4.73	26	-58	-32	544	3.98	8	-72	-26
	Cingulate-posterior											190	4.03	8	-44	30
	IFG-pars opercularis											234	3.58	42	16	14
2	IFG-pars triangularis											135	3.74	54	26	12
	Lingual Gyrus											22	3.43	12	-84	-12
	LOC						428	4.08	36	-78	40					
	MTG	1399	3.32	54	-42	-10	3975	4.87	50	-28	-10					
	SFG-dorsal/ paracingulate						204	3.42	4	28	44					
	STG											3875	4.33	66	-28	12
	Thalamus						226	4.18	8	-24	10	40	2.89	10	-28	14
	L LOC-inferior						139	3.03	-36	-58	18	3381	5.22	-50	-70	-20
	MTG-anterior	372	3.9	-50	-8	-20										
	MTG-posterior	3765	4.24	-54	-60	-8										
R	Precuneus-inferior															
	Precuneus-superior	741	3.8	-2	-70	46						866	4.54	2	-64	20
	SMG															
	SMG						437	3.38	-66	-52	16	2055	4.8	-54	-40	46
	Subcallosal gray															
	Angular Gyrus						714	3.34	62	-56	16	190	3.51	0	16	4

High Predictability		Scan 1					Scan 2					Scan 3				
IC	ROI	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z
3	Caudate-head	44	3.02	4	10	-4										
	Caudate-head	278	3.13	20	24	-8										
	Cingulate-anterior											147	3.72	2	16	34
	Insula-anterior											189	3.88	30	10	16
	Insula-middle											343	3.55	38	0	-14
	LOC-dorsolateral						355	2.97	36	-78	26					
	LOC-inferior	2282	3.9	52	-60	-6										
	MTG-anterior	159	3.49	54	-4	-30										
	MTG-posterior	1843	4.2	62	-40	4										
	PreCG-lateral	846	3.78	50	4	20	332	3.12	34	4	34					
	Precuneus-inferior	350	3.96	4	-58	34										
	STG-posterior											788	4.5	2	-64	28
	Angular Gyrus	594	3.37	-66	-44	10	86	3.82	-42	-50	22	6634	5.06	56	-28	16
	Anterior Insula											941	4.27	-52	-52	4
	Brainstem-midbrain											2368	4.45	-32	16	-16
	IFG-pars opercularis											174	4.48	-6	-28	-8
	MFG-anterior	2786	3.6	-42	24	10	196	3.01	-62	10	22					
	MFG-dorsolateral						410	4.09	-30	50	30					
	MFG-ventral											4494	4.26	-50	8	42
	MTG	592	3.91	-62	-38	-8	1026	4.46	-66	-20	-6	307	4.42	-26	50	24
R	MTG-posterior						57	3.38	-60	-58	-4	328	3.78	-68	-32	2
	Orbital Gyrus						467	4.05	-40	48	-18					
	PreCG-Inferior						35	2.97	-62	0	18					
	Precuneus-superior	27	3.3	2	-54	60						12	3.39	2	-62	60
	SFG-dorsomedial	784	3.7	-2	6	68						1184	3.71	2	22	40
	SFG-frontal pole											31	4.04	-20	58	-18
	SFG-ventromedial											91	3.46	-10	56	18
	SMG-inferior						366	3.45	-50	-56	34					
	Thalamus						306	3.63	-12	-22	-14	122	4.47	-6	-8	14
	Brainstem-midbrain											105	4.59	6	-30	-8

IC	ROI	Scan 1					Scan 2					Scan 3				
		voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z
4	L															
	Cingulate-posterior inferior											118	3.43	10	-46	12
	Fusiform-occipital division															
	IFG-pars opercularis	1507	3.7	54	20	-12	405	3.49	22	-68	-14					
	IFG-pars orbitalis						286	4.11	52	32	-16	672	4.02	60	14	-10
	Insula-anterior											823	3.78	38	12	6
	Insula-middle	29	3.19	36	2	12	34	2.95	40	4	6					
	Lingual Gyrus						420	4.06	16	-38	-4					
	LOC-dorsolateral															
	LOC-superior	227	3.49	14	-68	58						213	3.83	38	-60	52
	Precuneus-inferior											95	3.81	6	-62	60
	SFG-dorsomedial	474	3.73	10	14	60	496	4.12	8	16	56	40	3.47	20	-56	4
	SFG-ventromedial						36	3.35	4	46	28	197	3.98	10	18	60
	STG-temporal pole						1583	4.28	54	12	-8	92	3.09	8	50	26
	Thalamus															
	Amygdala											23	3.57	24	-26	14
	Angular Gyrus	3995	5.16	-48	-60	16						255	3.72	-20	-12	-14
	Brainstem-midbrain	371	3.68	-14	-32	74						3941	4.65	-36	-62	42
	Cerebellum-V	238	4.06	-24	-56	-18										
	Cingulate-anterior						64	3.13	-2	40	14					
	Cingulate-posterior											275	3.61	-16	-60	22
	Fusiform-temporal division	184	4.58	-28	-54	-10										
	Globus Palidus															
	IFG-pars opercularis	117	4.01	-56	16	2	138	3.87	-52	20	4	224	3.57	-24	-28	6
	IFG-pars orbitalis															
	MFG-dorsolateral						1280	3.95	-34	6	42	295	3.85	-50	42	-8
	MFG-frontal pole	158	4.32	-32	56	-16										
	MTG	1465	4.56	-54	-40	-14	1039	3.82	-50	-26	-8	1054	3.97	-58	-40	-6
	Orbital gyrus	794	4.57	-30	26	-22	614	4.15	-30	50	-10	449	3.96	-26	28	-14
	Parahippocampal gyrus	234	4.17	-24	-36	-10										

High Predictability		Scan 1						Scan 2						Scan 3					
IC	ROI	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z			
	Parahippocampal gyrus																		
	Precuneus-inferior						390	4.27	-10	-60	26								
	Precuneus-middle	941	4.54	-4	-70	32													
	Putamen	235	4.00	-20	6	12													
	SFG-dorsomedial	4160	4.91	-4	26	42													
	SFG-dorsomedial						1987	4.64	-2	24	48								
	SFG-dosolateral											5009	4.36	-4	32	40			
	SFG-frontal pole	532	4.01	-10	66	6						78	3.67	-6	60	-18			
	Thalamus	762	4.27	-16	-32	4													
R	Angular Gyrus	1836	5.09	52	-66	28	1659	5.02	56	-48	44	1946	4.11	56	-48	28			
	cerebellum-Crus II											398	3.84	28	-70	-42			
	Cerebellum-V						208	3.82	22	-44	-24								
	Cerebellum-Vermis	37	3.86	4	-58	-10													
	Cerebellum-VI	217	5.07	24	-48	-24													
	Cingulate-anterior						44	3.21	2	40	14								
	Lingual/Cerebellum	1032	4.38	26	-62	-4	995	4.05	38	-42	-14	482	4.07	20	-80	-16			
	Lingual Gyrus	137	3.21	16	-70	-12						355	3.85	30	-46	-4			
	MTG											215	3.45	64	-18	-26			
	Parietal operculum	824	4.84	52	-8	14	231	3.50	42	-30	26	1251	3.82	64	2	4			
	PostCG-dorsal	782	4.53	26	-34	62	298	4.05	22	-34	68	652	4.31	24	-38	66			
	PreCG-dorsal						26	3.23	10	-12	66								
	Precuneus-inferior	54	4.27	2	-80	34													
	SFG-dorsomedial	541	4.49	2	28	44	346	4.91	2	24	48	309	3.86	6	32	48			
	SFG-ventromedial											77	3.87	2	56	14			
	STG	948	5.05	64	-16	0	1183	3.94	56	-22	8								
5	L Cerebellum-crus I						111	3.87	-48	-52	-28								
	Cingulate-anterior	230	4.00	0	-2	28	154	3.34	-2	-12	30	343	4.39	-4	-6	28			
	Cingulate-posterior	654	3.94	-10	-46	12	117	3.78	-2	-32	24	94	4.66	-6	-40	18			
	Cingulate-posterior inferior											54	4.92	-4	-42	18			

High Predictability																
IC	ROI	Scan 1						Scan 2						Scan 3		
		voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z
	Cuneate gyrus						189	3.21	-2	-76	10	230	3.97	-2	-86	14
	Fusiform-Temporal division											1576	4.12	-42	-36	-22
	ITG-posterior	986	4.26	-44	-52	-4	1207	3.57	-50	-80	-6					
	LOC-superior						384	3.98	-34	-74	44	4508	4.87	-28	-62	48
	MFG-dorsolateral											400	4.68	-36	36	22
	PreCG-dorsolateral	722	4.36	-30	-2	32	611	3.65	-50	0	36	511	4.18	-48	6	44
	SFG--ventromedial	290	4.67	-4	58	-8										
	SFG-frontal pole						430	3.92	0	56	-8	444	4.11	-6	52	-10
	SMG	3661	4.63	-56	-40	54										
	SPL						3440	3.90	-24	-60	56					
R	Angular Gyrus											2752	4.19	44	-42	54
	Cerebellum-crus I	292	4.56	34	-68	-36										
	Cingulate-anterior											328	3.778	6	-16	36
													03			
	Cingulate-posterior	821	3.69	10	-46	10										
	Cingulate-posterior inferior						241	3.72	10	-34	32					
	Cuneate gyrus						133	3.25	2	-76	10	98	4.02	2	-86	14
	IFG-pars opercularis											315	4.91	46	8	6
	IFG-pars triangularis	304	4.09	44	38	6	524	3.28	50	40	4					
	ITG	315	5.10	52	-56	-26	241	3.44	54	-50	-28	587	3.75	52	-62	-30
	LOC-superior	1018	4.69	30	-58	32	848	4.13	36	-66	46					
	SFG-frontal pole	242	3.82	6	58	-6	376	3.89	18	62	-2	474	4.56	12	52	-4
	SMG	950	4.45	54	-42	48	1301	3.50	62	-38	48					
	STG-temporal pole						39	3.05	52	12	-6	131	3.46	46	2	-10
Low Predictability																
IC	ROI	Scan 1						Scan 2						Scan 3		
		voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	vox	Peak t	X	Y	Z
1	L Cingulate-anterior-ventral	135	3.43	4	44	14										
	Heschl's Gyrus											433	3.42	-34	-26	-2
	PostCG-dorsolateral	251	4.37	-48	-24	56										

Low Predictability		Scan 1					Scan 2					Scan 3				
IC	ROI	voxels	Peak t	X	Y	Z	voxels	Peak t	X	Y	Z	vox	Peak t	X	Y	Z
R	STG	4129	3.9	-48	-24	0	1282	3.35	-58	-4	-14	802	3.68	-62	-20	14
	Cingulate-anterior-ventral	413	3.48	6	44	4										
	MTG-posterior											1394	3.63	36	-24	2
	STG	3820	3.95	44	-38	0	1948	3.57	50	-40	8					
2 L	MFG-dorsolateral						237	4.2	-32	10	34	488	3.33	-32	16	32
	PreCG-dorsolateral											270	3.7	-30	-8	36
	SFG-dorsolateral											24	3.49	-22	18	44
	SPL-inferior						377	3.8	-28	-46	40					
R	Anterior Insula						782	3.51	28	42	-4					
	Caudate						143	3.68	10	8	10					
	Fusiform Gyrus											46	3.4	46	-58	-24
	IFG-pars opercularis						71	3.08	30	14	22					
	IFG-pars triangularis											1282	3.92	52	20	-4
	LOC						389	3.84	26	-64	42					
	MFG-dorsolateral						294	3.96	28	14	46					
	MFG-ventrolateral						1209	4.01	50	46	18					
	MTG-posterior						243	3.43	44	-48	4	194	3.77	46	-50	-4
	SFG-dorsomedial						176	3.14	4	22	38					

Voxels are 2x2x2mm. Coordinates refer to MNI space. IFG-Inferior Frontal Gyrus, ITG-Inferior Temporal Gyrus, LOC-Lateral Occipital Gyrus, MFG-Middle Frontal Gyrus, MTG-Middle Temporal Gyrus, PostCG-Postcentral Gyrus, PreCG-Precentral Gyrus, SFG-Superior Frontal Gyrus, SMG-Supramarginal Gyrus, SPL-Superior Parietal Lobule, STG-Superior Temporal Gyrus