Syntax processing in language and mathematical formulas

Martin Perez-Guevara
Supervisor: Christophe Pallier

ABSTRACT: Diverse studies in neuroimaging have been conducted to characterize the network of brain areas related to the processing of natural language and more specifically to study effects of syntactic manipulation in language. Nonetheless less neuroimaging studies exist exploring the network of brain ares related to the processing of mathematical stimuli and the topic of syntactic manipulations in mathematics is relatively unexplored. In this research project some datasets from previous neuroimaging experimental designs will be analyzed and furthermore new experimental designs will proposed, conducted and analyzed to contribute to the existing evidence behind the characterization of the network of brain areas related to the processing of mathematical stimuli and to more specific syntactic manipulations in mathematics.

◆

1. PRELIMINARY SUMMARY OF THE INTERNSHIP EXPERIENCE, WORK CONDUCTED AND CHALLENGES ENCOUNTERED.

Before developing the main content of this report I would like to summarize all the different efforts that comprised my research experience during my internship in the Unicog lab in NeuroSpin under the supervision of Christophe Pallier.

As I found out, working in the neuro-imaging field can be extremely challenging, since not only one has to become an expert on an specific topic of cognitive neuroscience to propose adequate experimental designs but moreover one has to become well acquainted with all the principles behind the machinery and technology employed to understand all the data acquisition, preprocessing and processing steps. These are not at all trivial in the case of MRI, which in itself is a modern technology still sensitive to advances in research and development. Furthermore a high level of mathematical, statistical and programming skills is required to be able to successfully conduct experiments and their analysis with certain level of independence. So far, I think this have been the most demanding and rich experience of work and research that I had in my life.

During my research internship I accomplished many feats that will comprise what is presented in this report. Although it will seem in some cases that what I present still lacks many steps of analysis, I want to emphasize that neuroimaging experiments are normally expected to last 1 year, from the design of an experiment to its complete, although not necessarily exhaustive, analysis. In my case not only I approached a complete new field during the past 6 months but on top of that I was able to effectively design experiments and contribute to the implemented methods and programming pipeline of the laboratory for data analysis. My achievements could be summarized as:

1. Proposing effective and varied experimental designs to contribute to the state of the art knowledge of the given questions in this research project.
2. Conducting a pilot, based on the proposed experimental designs.
3. Conducting the acquisition and partial analysis (16 subjects out of 20 scanned to this day) of a complete experiment setup based on the proposed experimental designs, refined after the conduction of the pilot study. (The pilot study is not developed in this report for its methodological overlap with the complete experiment)
4. Developing an additional programmatic pipeline of analysis that could be reused in the laboratory for the general analysis of experimental designs that employs new tools developed at the laboratory and fulfills other purposes.
5. Realizing all the preliminary individual preprocessing and processing and group analysis of a previously developed experimental design, for which data was acquired, but analysis was pending. Taking advantage of the programmed pipeline.

2. INTRODUCTION

Language is one of the most studied cognitive skills, since it sustains cooperation and knowledge accumulation in society. The way in which we process language can radically impact mutual understanding during communication and impair or improve learning. For this reason discovering how we process structure in language, as one of its fundamental aspects for production and comprehension, is a very important topic of research.

In the first place, we know that the network of brain areas involved in parsing sentences has been identified by Pallier, Devauchelle & Dehaene (2011) who observed increased activations in areas as a function of the size of syntactic constituents. In the case of mathematics, Friedrich & Frederici (2009) have presented hierarchical or flat logical formulas to subjects and, from the contrast between the two, concluded that maths and language rely on different networks. Moreover, Maruyama et al. (2012) have come to a similar conclusion by presenting subjects with more or less unstructured arithmetic formulas. The fact that mathematics and natural language might be processed by different networks in the brain becomes then an important topic of research to understand how the brain manipulates symbolic information.

Moreover, neuroimaging is one of the most important tools in the advancement of the field of cognitive neuroscience. A great quantity of studies employing FMRI and MEG have been conducted to understand the brain areas related to the processing of language, as reviewed by Friederici (2011), Price (2012) and Tyler et al. (2008). Also, more specifically, syntactic manipulations have been studied, as reviewed by Kaan et al. (2002) and Grodzinsky et al. (2006). But, in contrast, only two studies on syntax processing of mathematical stimuli exist (Friedrich & Frederici 2009 and Maruyama et al. 2012), since most of neuroimaging studies in mathematics explore the topic of numerosity. In this research project we will particularly explore Bold-FMRI studies to provide supporting evidence for the identification of the brain areas related to the general processing of syntax in mathematics and to more specific syntactic effects.

Due to the latest mentioned studies several important questions arise regarding the encoding, complexity and processing of mathematical syntactic structures in the brain. How much overlap should we expect between the brain regions responsible for language and mathematical processing, specially when we consider syntactic manipulations? Can we find areas that would respond exclusively to syntactic manipulations in mathematical formulas? How mathematical syntactic manipulations are processed in the brain, can we show areas that respond to increasing complexity? To be able to approach some of these questions diverse experimental designs based on neuroimaging are explored in this research project.

3. THEORETICAL BACKGROUND

2.1 Cognitive Neuroscience in Language (problems related to syntax)

Although written sentences in natural languages or mathematical expressions are typically expressed by linear series of symbols, there is some evidence that they are internally represented by hierarchical tree structures. In the case of sentences, classic arguments for the existence of such structures come from Linguistics (Chomsky, 1957; Jackendorf , 2008; Bake & al, 2001). They are normally analyzed under a generative grammar framework and most commonly related to context free grammars and their corresponding parsing strategies (Jurafsky & Martin, 2000). In the case for mathematical expressions the structure can, and sometimes needs to, be cued by parentheses and there is also some cueing of grouping of words by prosody if we consider spoken utterances. Which makes the tree representation of mathematical expressions even more plausible. In Figure 1 we show an example of what parsing of a mathematical expression would look like.

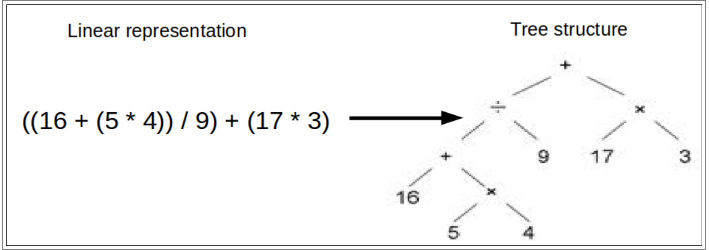


Figure 1. Example of parsing. (tree structure extraction from a linear representation)

Evidence for the cognitive processes behind syntactic structures can be tracked down to priming studies demonstrating how people tend to reuse syntactic constructions when they are previously exposed to them, (Branigan 2007 and Noppeney & Price 2004) reported that syntactic priming caused a repetition suppression effect in a region of the anterior left temporal lobe. Other brain areas have also been related to syntactic processing, like the left inferior frontal gyrus, including Broca's area (Tettamanti al, 2002; Musso et al, 2003; Ben-Shachar et al , 2003; Ben-Shachar, Palti, and Grodzinsky, 2004).

Interestingly Tettamanti et al (2002) and Musso et al (2003) taught grammatical and ungrammatical rules to participants during FMRI scanning. The grammatical stimuli depended on hierarchical structures while the ungrammatical one only had sequential properties. It turned out to be the case that subjects learned the rules behind both types of stimuli but the grammatical ones show more activation in the left inferior frontal gyrus. On the other hand, evidence for the participation of Broca's area in the processing of hierarchical linguistic structures come from studies employing artificial grammars (Friederici et al, 2006), in which the subjects were trained in two languages, one with embedded constituents and the other only with serial structure. It turned out that only participants that learned the embedded constituents structures showed increased activation in Broca's area.

2.3 Cognitive Neuroscience in Mathematics (Problems related to syntax)

In the case of mathematical formulas, little research exists about their parsing and the brain areas involved. Jansen, Mariott and Yelland (2003) showed that well-formed formulas are better memorized than ill-formed ones. Also they showed in later research (Jansen, Mariott and Yelland, 2007) that the latencies of fixation increase at the end of mathematical constituents, in a similar fashion to text. This suggested that mathematical formulas rely on similar mechanisms as language even though it has other forms of presentation, sometimes two dimensional.

Moreover, Friedrich & Friederici (2009), who studied complexity in mathematical expressions in a similar way to studies that compare sentences and lists of words, were surprised to find limited activation in frontal areas, since their earlier work on language revealed more posterior regions. Furthermore, other behavioral studies exist that have shown more specific syntactic effects in mathematical expressions. For example Landy and Goldstone (2007a, 2007b) showed how subjects expressed and were influenced in different tasks by spatial effects in the production and perception of mathematical formulas. Nonetheless, as mentioned before the behavioral evidence behind parsing in mathematical expressions is still limited in comparison to language and lacking a neuroimaging counterpart.

2.2 About Bold-Fmri

Magneto resonance imaging (MRI) have become an important technology in the medical and research domain. One could say it is at the heart of cognitive neuroscience and of the integration of multiple disciplines of study. Furthermore diverse machinery and techniques exist for the implementation of Functional MRI (FMRI). One of which is the Blood Oxygen Level Dependent FMRI (BOLD-FMRI).

BOLD-FMRI is a relatively recent technology, with its first studies showing how simple sensory stimulation could lead to changes in a blood oxygenation level dependent contrast, as in Ogawa et al. (1992). It exploits the fact that “ferrous iron on the heme of deoxyhemoglobin is paramagnetic, but diamagnetic in oxyhemoglobin” Chen et al. (1996). Moreover the hemodynamic response of the brain to a stimuli can be quite complex as explained by Logothetis (2003). Nonetheless Boynton et al. (1996) showed that under many conditions a regression linear model employing a double gamma function as basis can capture quite well the hemodynamic response function (HRF) and that the perception of two stimuli add quite linearly as long as they are separated enough in time, more than 2 seconds according to Buckner (1998). In Figure 2 we can see an example of the model fitted by Boynton.

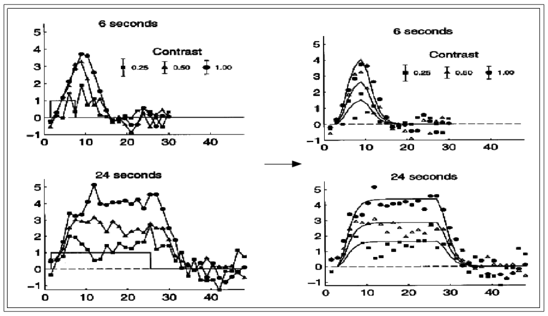


FIGURE 2. HRF MODEL FIT, TAKEN FROM BOYNTON (1996)

The process behind an experimental design, data acquisition, preprocessing and analysis when employing Bold-Fmri can be quite complex and even nowadays not completely mastered as explained by Strother (2006) in Figure 3. Moreover one has to be careful in the interpretation of changes in the Bold signal as explained by Logothetis (2008). This leads to the fact that every step around experimentation employing the technique of Bold-Fmri have to be greatly controlled and understood. Experimental design in particular is quite sensitive to the peculiarities of the technique due to the nonlinearities that might arise if time between stimuli is not appropriately controlled and the statistical properties of the models used to estimate the HRF and detect changes in the bold-signal, which efficiency is greatly influenced by the sequences of stimuli presentation as explained by Henson (2007)

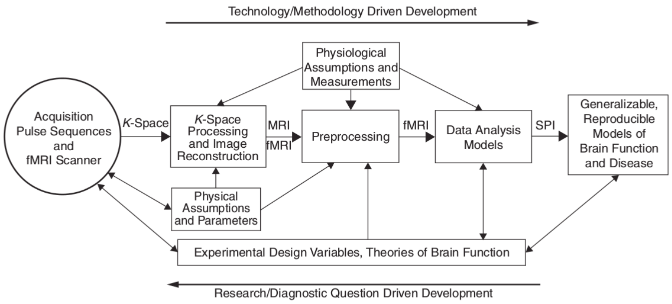


FIGURE 3. GENERAL DIAGRAM OF FMRI EXPERIMENTAL PROCESS. TAKEN FROM STROTHER (2006)

Once reconstruction is performed from the data recollected with the pulse sequences of the FMRI scanner, it is necessary to realize several manipulations in the data called preprocessing before any analysis can take place. There are in general 4 steps in the preprocessing. Since acquisition is really performed by slices but we are interested in analyzing the brain as a complete volume, first a slice timing algorithm have to be implemented to extrapolate to a common time all slices corresponding to a theoretical volume. Then due to possible movements of the subject in the scanner a coregistration algorithm have to be implemented, such that we can analyze each voxel of every volume as having a history of bold-signal. In the third place it is necessary to project the brain activations of all subjects to a common normalized space where they can be compared and put together for a group analysis. Finally, taking also into account that there are important inter-subject variabilities as shown by Handwerker (2004), smoothing algorithms are normally applied to the normalized data of each subject to facilitate the identification of common regions of activation and diminish the influence of noise in the estimations. More details can be read from Lindquist (2008).

After preprocessing of the data is complete, the statistical analysis relating the stimuli presented for a given experimental design and the HRF activations can be performed. As explained by Lindquist (2008) and based on the work presented by Boynton (1996), the classic analysis involves the application of a generalized linear model to estimate one parameter per stimulation condition defined in a experimental design. Experimental conclusions of the relationship and effects of the stimuli will be drawn from the contrasts of the estimators of the mean hemodynamic response of the brain in each of the voxels considered for each volume of acquisition. We present a graphical example of the GLM estimation for two types of stimuli given a bold-signal in a specified voxel.

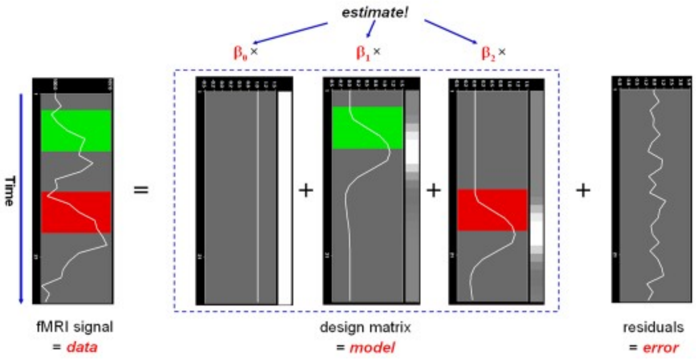


FIGURE 4. EXAMPLE OF HRF ESTIMATION OF TWO CONDITIONS WITH A GENERALIZED LINEAR MODEL (GLM). TAKEN FROM THE BRAINVOYAGER USER GUIDE.

Furthermore, the statistical complexity of the analysis increases when we consider the problem of multiple comparison that we get from thousands of parallel estimations, considering that a volume of acquisition for a resolution of 1.5mmx1.5mmx1.5mm can have around 500.000 voxels. Due to this, diverse methods employing random field theory or other mathematical theories have been developed and FWE or FDR corrected p-values are considered for the confirmation of effects, as proposed by Worsley et al. (1996).

4. GENERAL RESEARCH PROBLEM AND APPROACH

There are three main state of the art questions that I am approaching in this research project.

1. What are the brain areas that respond to mathematical formulas processing?

2. What are the brain areas that respond more specifically to syntactic manipulations in mathematical formulas? Are there areas that reflect automatic parsing?

3. Can we find areas that respond to complexity in number of constituents (number of valid operations) in mathematical formulas?

The first and third questions have been approached by recent experiments in the lab. Unpublished results on general research on the processing of mathematical concepts written on words and mathematical formulas seem to recruit additional areas that are not employed by natural language. It seems that there are important non overlapping regions.

Moreover Maruyama et al. (2012) In the paper “The cortical representation of simple mathematical expressions” published in 2012 reported results from experiments at the lab that supported the notion of non overlapping areas for mathematical processing and indicated areas that responded to complexity in the number of constituents in formulas. Nonetheless such effects could be influenced by the employment of a demanding memory task during the presentation of the designed formulas. Moreover automatic parsing areas could not be identified since the parsing of formulas is forced by the task.

The second question on the other hand has not been directly tackled by current experiments, specially because of the complications behind exploring syntactic manipulations in mathematical formulas. This is due to the fact that the possible syntactic manipulations are very restricted if one wants to keep the grammaticality of the formulas and also the number of constituents. Moreover possible changes in syntactic structure are correlated with number of characters and visual effects due to the introduction of parentheses.

To address the depicted questions, two experimental designs are proposed and analyzed alongside a simple language localizer, which is an experimental design focused on identifying the network of brain areas processing language, and a previous experiment conducted at the laboratory based on the problem of syntactic composition in language.

The first experimental design proposed is mainly focused on question 3. It consisted on employing as stimuli expanding embedded trees similar to those use in Maruyama's experiment, that consisted on formulas that had the same length with different number of constituents embedded in non grammatical character sequences. But in this case we completely got rid of the memory task, so the stimuli is simply presented for very quick intervals of time (200ms) that are even faster than possible eye saccades. The task in this case would be a simple detection instruction to ensure attention but without forcing the parsing of the stimulus.

We would expect to confirm decreasing complexity effects reported in Maruyama's paper, observe the network of brain activations related only with the appreciation of the mathematical formulas and moreover be confident about the relationship between these brain areas and automatic processes since parsing was not forced. Moreover we extended Maruyama's stimuli to also include formulas with symbols since the original stimuli only contained digits, so that we could differentiate areas that responded to complexity for symbolic and numeric representations in formulas.

The second experimental design proposed is mainly focused on question 2. I proposed to implement a syntactic supra-conscious priming design on mathematical formulas. The idea is that by having differences in the shape of formulas but not in the syntactic tree structure we could isolate the effect on brain activations of syntactic manipulations. Due to the repetition suppression effect that can be seen in the hemodynamic response function we expect to test the hypothesis that an area involved in automatic parsing processing would show more activation for pairs of formulas with different tree structure than with the same tree structure.

Moreover we include non formula stimuli with the same characters and similar visual shape manipulations as the ones induced in the real formulas to control for areas that might be responding to mere visual priming effect. Again on this experiment, as in the first one, the task during each trial was a simple detection task that would need interfere with formula parsing and would not force parsing, but would just ensure attention to the stimuli. In addition we considered two levels of formula complexity, expecting that a simple contrast should confirm some areas as revealed in Maruyama's experiment.

5. EXPERIMENTS IN MATHEMATICS (SHARED METHODS)

During my internship I proposed three experimental designs to tackle the previously presented research problem on the syntax and complexity of mathematical formulas. In this section I provide general methodological details concerning all experimental designs since they were implemented together for all subjects in one scanning session of around 1 hour and 6 minutes.

First a pilot was run for 2 subjects. This gave interesting preliminary results at the subject level that justified running a complete experiment with 20 subjects. Part of such results was a possible reconfirmation of expected areas of activation due to constituent complexity in mathematical formulas as presented by Maruyama et al. (2012). Details on the pilot will be omitted in this research project report for being quite similar to the final design before some optimizations in the stimuli and its presentation.

20 was the number of subjects chosen since its a standard rule of thumb in neuroimaging experiments to look at group results. It is expected that a sample of such size already reveals some significant effects at corrected p values. Up to date, 16 subjects have been scanned. So this report will present only preliminary results of the group analysis expecting that adding 4 more subjects to the sample would not dramatically change the presented results and analysis. Nonetheless there are still many steps on the analysis and cleaning of data that for time reasons have not been done and need to be conducted and this could have an important impact in the final results in contrast to the ones presented here.

5.1 Subjects

16 native French speakers participated in the FMRI experiment (9 females 56,25%; age range 18.6-26.9 years, mean 23.4, SD 2.4). All subjects had a scientific high-school background from French universities (Bac S) and were right-handed with a laterality quotient (LQ) of at least 40 (mean 68.1, SD 14.2), as measured by the Edinburgh Handedness Inventory (Oldfield, 1971) . The experiment was sponsored by the language neuroimaging unit of Unicog lab in NeuroSpin under the supervision of Christophe Pallier as director of the language neuroimaging team. Moreover the experiment received ethical approval by the regional ethical committee (Comité de Protection des Personnes, hôpital de Bicêtre).

5.2 Session Design (Presentation of experiments)

The session was designed such that the three experimental designs would be included in one session along the acquisition of the anatomical image employed as reference for preprocessing of the FMRI data . Experiments where arranged such that acquisition would normally occur in a total time of 1 hour, 6 minutes and 37 seconds in the following order:

1. Anatomical T1 (7 min 46 sec, 1 volume)
2. Run 1 of Expanding Trees (6 min 32 sec, 246 volumes)
3. Run 2 of Expanding Trees (6 min 32 sec, 246 volumes)
4. Run 3 of Expanding Trees (6 min 32 sec, 246 volumes)
5. Run 4 of Expanding Trees (6 min 32 sec, 246 volumes)
6. Run 1 of Priming Trees (6 min 32 sec, 246 volumes)
7. Run 2 of Priming Trees (6 min 32 sec, 246 volumes)
8. Run 3 of Priming Trees (6 min 32 sec, 246 volumes)
9. Run 4 of Priming Trees (6 min 32 sec, 246 volumes)
10. Language Localizer (6 min 35 sec, 248 volumes)

Only the anatomical scan has a different set of parameters than the rest of the runs. All experimental designs had the same scanning parameters except for the language localizer having an additional 2 volumes of acquisition. Details on the scanning parameters of acquisition for the anatomy and rest of the scans are detailed in section 6.5.

5.3 Stimuli presentation

Stimuli in all experimental designs were presented using a fixed-point Inconsolata font. Projected on a translucent screen with a digital-light-processing projector (PT-D7700E, panasonic; frame rate: 60 Hz, resolution of 1024 x 768), with a viewing distance of a 89 cm. All experimental designs were implemented with python scripts mainly exploiting the capabilities of the Expyriment python library (Krause, 2014). On top of which careful programming had to be implemented to ensure the correct timi.ng of stimuli presentation.

5.3.1 Notes on the generation of the sequence of stimuli presentation.

As part of the methodology of experimental design, I employed Wager's genetic algorithms (Wager and Nichols, 2003) to design the sequences of presentation of stimuli based on a A-optimality criterion. We show in Figure 5 examples of how the timing and sequence of stimuli presentation can have an impact in the theoretical predicted bold-signal. Which is further explained by Henson (2007).

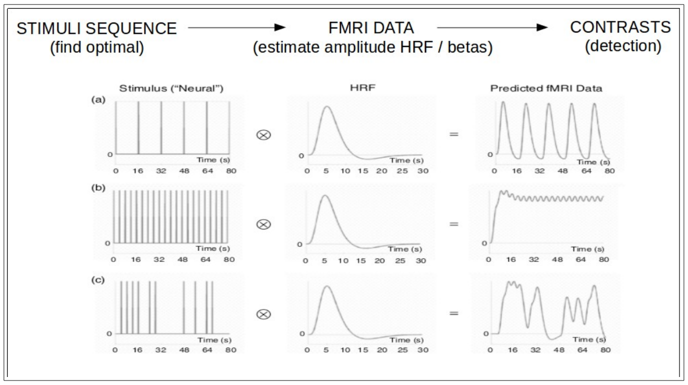


Figure 5. Example of influence of the timing and sequence of stimuli in the predicted bold-signal.

Under this framework, we gave a higher weight to detection over estimation for the main expected contrasts, since direct HRF estimation would not be the main goal and is additionally improved by jittering the onset.

The contrasts considered for the experimental design Expanding Trees were: Linear effect on the number of valid operations for the 4 levels considered in the mathematical expressions designed and the effect of comparing mathematical expressions with characters and with digits.

On the other hand the contrasts considered for the experimental design Priming Trees were: Comparing structured and unstructured mathematical expressions; Comparing pairs of mathematical expressions with a syntactic difference with pairs with the same tree structure, under the interaction with structured over unstructured expressions; Comparing mathematical expressions with two operators with mathematical expressions with one operator, under the interaction with structured over unstructured expressions.

5.4 Behavioral data analyses

In the case of the language localizer the task would simply be to look attentively at the stimulus. On the other hand, the task employed in both mathematical experimental designs (Expanding Trees and Priming Trees) was simply to detect an "@" symbol inside the presented mathematical stimulus. The idea behind employing such a simple detection task was to secure attention of the stimulus without interfering with its processing in the cases when there should be no response to the task (stimulus without an "@" symbol inside).

To express the detection, the subject would press a button with their right hand only when detecting the "@" symbol. Any stimulus modified for the task would just be employed as a representative category for a button press in case the subject responded correctly to the task. Any stimulus not modified for the task for which a false positive detection (button press) occurred would be discarded, as well as any task modified stimulus for which there would not be a detection (lack of button press).

Due to this design the behavioral analysis is quite simple and just related with the assessment of the attention given to the stimulus by the subjects. Basically we will consider for each experimental design the correct responses to the task as a measure of attention and the amount of false positives as a measure of the amount of clean stimulus negatively affected by the task. In general we will see in posterior reports of the experimental designs that most subjects had very high accuracies and extremely low false positive responses. Which supports the success and usefulness of the task.

5.5 FMRI acquisition and analyses

All MRI data was acquired on a 3 Tesla MR system (Siemens TrioTim Syngo). For each subject, anatomical images were obtained using a sequence of magnetization prepared rapid gradient echo (MPRAGE) with 160 slices with a voxel size of 1 × 1 × 1.1 mm, covering the entire brain (TR = 2.3 s, TE = 2.98 ms, flip angle = 9°, no gap).

FMRI data was acquired as T2*-weighted echo-planar image (Multi-Band EPI C2P from Minnesota University) volumes. The MultiBand EPI consisted on the parallel acquisition of 4 slices at a time, reconstructed by a parallel imaging reconstruction algorithm (Chaâri et al, 2010). Eighty transverse slices covering the whole brain were obtained with a TR of 1.5 s and a voxel size of 1.5 × 1.5 × 1.5 mm (TE = 32 ms, flip angle = 70°, no gap). For all runs 246 volumes were acquired except for the language localizer, for which 248 volumes were acquired. Accurate timing of the stimuli relative to FMRI acquisition was achieved with an electronic trigger at the beginning of each run.

All data was processed with FSL (Smith et al, 2004) and SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). In the first place, anatomical images were normalized to the standard brain template defined by the Montreal Neurological 152-brains average. Regarding functional images, slice timing corrections were done to have an estimated equivalent time of acquisition for all slices in a volume. Moreover, due to the parallel acquisition of slices, the slice timing procedure was run on FSL software since it allows for the specification of the simultaneously acquired slices. Then the functional images were realigned to correct for head movements employing a rigid body transformation. Such that the first image of all runs are aligned with the first image of the first run and all other images are aligned with the first image of their respective run. The rigid body transformations applied to each volume were saved so they could later be employed as regression parameters on the Generalized Linear Model employed for analysis. Afterwards, the functional images were spatially normalized using the parameters obtained in the normalization of the anatomical images, resampled with a voxel size of 1.5 × 1.5 × 1.5 mm, and smoothed with a 4.5 mm Gaussian kernel.

A general linear model was created, where experimental effects at each voxel were estimated using a multi-session design matrix. All trial types were modeled by the canonical hemodynamic response function, its first-order time derivative, its first-order space derivative and the 6 individual motion parameters extracted from the rigid body transformation to capture remaining signal variations due to head movements. The model also included high-pass filtering above 1/128 Hz. Individually estimated contrasts on the BOLD responses, smoothed with an 8 mm Gaussian kernel, were then entered into a random-effect group analysis. The mentioned preliminary group analysis consisted on a simple T-Test implemented on all the contrasts computed for each subject for each experimental design. Furthermore in some cases we will present overlaps of the activation of brain areas related to different contrasts, which is possible by employing the xjView toolbox (http://www.alivelearn.net/xjview).

Moreover, for all the preprocessing and processing of data for individual and group analysis in all the experimental designs, an automatized pipeline was developed in Bash Scripting and Matlab code to be run on the Linux O.S. The pipeline mainly exploits standard algorithms contained in diverse packages like FSL (Smith et al, 2004), SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) and PyPreprocess (http://github.com/neurospin/pypreprocess). This common pipeline was developed with the aim of replacing easily Matlab code in the future by python implementations of the processing and preprocessing steps. Also it would allow easy update of the analysis as subjects are scanned and would save time by allowing all experimental designs to be analyzed under a common pipeline. Details on the pipeline developed can be seen in Appendix A.

5.6 FMRI analysis important clarifications

It is important to note that data from 4 subjects is missing in the group analysis and although it should not dramatically change the general picture of the results it can have an important impact in the corrected significance. For this reason we might also explore results that are significant for uncorrected p-values of 0.001 and for which Family Wise Error corrected p-values (FWE) are at least below 0.5. Considering both peak levels and clusters of activation. Any peaks of activation with high FWE p-values (above 0.5) will tend to be disregarded as noise in this report.

Moreover, due to time constraints, for this report we have not been able to perform all the manual quality checks on the individual analyses, which could lead to extra preprocessing in the case of some subjects because of specific artifacts or could even lead to dropping some experimental runs due to behavioral problems or extreme artifacts. This could also have an important impact in the results.

Any results for which FWE corrected p-values achieve significance with values below 0.05 will be emphasized. Nonetheless it is important to notice that even such significant values have to be reported with caution due to the sensitivity of FMRI data to noise artifacts. Corrected significant evidence appearing simultaneously in diverse experimental designs would be the ideal scenario to report results or make hypothesis with a high confidence.

6. EXPERIMENTS IN MATHEMATICS (LANGUAGE LOCALIZER)

The language localizer is a simple block design aiming to identify brain areas responsible for language processing to be able to compare them with the brain areas that would respond to mathematical stimuli. This experimental design was previously developed at the laboratory and successfully employed in other experiments. Moreover the identification of the brain areas related to language in each subject can be employed as a sanity check of the FMRI acquisition, since these areas have already been well depicted in the literature.

6.1 Experimental design

6.1.1 Stimuli

The stimuli consisted on blocks of sequences of words that form a phrase and blocks of sequences of non words. These blocks were presented in an alternated fashion and the main idea is that we could extract the brain areas processing language from the contrast of this block types. The presented text in the screen comprised 0.72° of vertical visual angle and a maximum of 5.8° of horizontal visual angle, considering that words with different length are presented and the largest word contains 14 characters.

6.1.2 Task

As explained in the previous shared methods section, the task was simply to look attentively at the text presented on the screen.

6.1.3 Trial design

Each trial mainly consisted on presenting one of the blocks designed (consisting on phrases or non words). Each block contains three sequences of text units, the first sequence made of 9 units, the second of 10 units and the last of 9 units. A fixation cross would be presented before the presentation of each text sequence, for 500 ms, followed by a blank screen for 500 ms. Then for the presentation of a sequence, each text unit would be presented for 200 ms. Between the presentation of the three different sequences a blank screen would be presented for 600 ms. At the end of the presentation of the three sequences a blank screen would be presented for 7 seconds waiting for the next trial (the next block). There were 4 runs of acquisition and in each of them 90 trials were presented. In Figure 6 we can appreciate a diagram of the trial design.

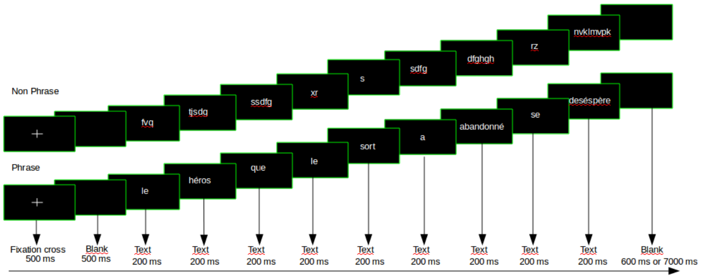


FIGURE 6. TRIAL DESIGN OF THE LANGUAGE LOCALIZER

6.2 Fmri Results

6.2.1 Questions of interest and main contrasts to look at.

In this experimental design we just want to look at the language related brain areas. Basically we will look at the contrast between the phrases blocks and the non words blocks. This would also serve as an additional sanity check for all the FMRI acquisitions.

Questions list:

1. Which are the language related brain areas?

Corresponding main contrasts lists:

1. Phrase Blocks over Non Words Blocks

6.2.2 Group Results

1. We find extended activation on the left temporal and left precentral areas. Moreover these activations have peaks and sub-clusters significant at FWE p-values < 0.05. Details on the activation maps and the cluster statistics can be seen in the Appendix B.

6.3 Discussion

As expected we clearly see activations of the well known language areas as summarized by Friederici (2011). We see the activation of Broca's area in the inferior frontal gyrus (IFG), Wernicke's area in the superior temporal gyrus (STG), as well as parts of the middle temporal gyrus (MTG) and the inferior parietal and angular gyrus in the parietal lobe. Moreover these activations give us the required sanity check to look into the FMRI acquisition data of the other experimental designs.

7. EXPERIMENTS IN MATHEMATICS (EXPANDING TREES)

The expanding trees experimental design is inspired on the work published by Maruyama et al. (2012), that was previously developed by the language neuroimaging team. It consists on employing as stimuli expanding embedded trees as those used in Maruyama's experiment. But in this case the main idea was to explore if the same areas and effect observed in Maruyama would remain if we simplified the task, taking away the memory effort. The complexity and other effects reported in the paper could actually reflect memory processes instead of syntactic processing of the formulas. Moreover by having only a very quick presentation of the formulas we could explore possible automatic parsing processes.

7.1 Experimental design

7.1.1 Stimuli

So our stimuli had the same structure as Maruyama. With the addition of symbols to the analysis of the formulas and not only digits. We would expect that adding symbols should not change the brain activation effects related to parsing.

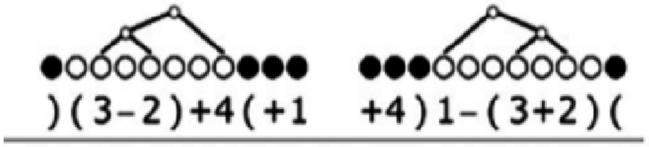
Employing the software Wolfram Mathematica I extracted the structure templates of the Maruyama experiment and reproduced them with new formulas with digits 1,2,3 and 4 and symbols a,b,c and d. The formulas consist of eleven symbols always with 4 digits or letters, two or one "+" sign, correspondingly one or two "-" sign, and two pairs of round parentheses. The ordering of the symbols determine four levels of structure complexity, were the first level (0 nested levels of calculation) is completely unstructured and the last level (3 nested levels of calculation) is completely structured in a tree with left or right branching (considering the position of the most inner parentheses). An example of the stimuli for different levels with digits and symbols, as well as a small diagram showing the possible branching difference is shown in Figure 7. More details in the original template generation can be revised in Maruyama et al. (2012). Also more details on the stimuli categories and proportion in each run can be seen in Appendix C, in Figure C.1 a complete diagram of branching for digit formulas can be seen, while in Table C.1 a complete example with all stimuli and their presentation proportion in each run is given.

As an additional modification of the stimuli, we exchanged one symbol in a formula, uniformly selected from the formula categories, by an "@" symbol. Since this set of modified formulas would be employed for the task.

As a last remark, the presented formulas in the screen comprised 0.72° of vertical visual angle and 4.44° of horizontal visual angle. Considering that formulas contain 11 characters.

7.1.2 Task

As explained in the previous shared methods section. The task is simply to detect an "@" symbol inside the presented mathematical stimulus. In which case, the subject had to press a button with the right hand as soon as possible.



Category	Ex. numbers	Ex. symbols
Level 3	((2+3)-4)+1	((a+b)-d)+c
Level 2	)(4+3)-1(2-	)(a+d)-c(b-
Level 1	2)1-4((+)3-)	)d+b(-c-(a
Level 0	4)3-+(1()2-	+)-c)d+(a(b
Task	3+(@2)4)(1	b(-+d)c@)-a

FIGURE 7. STIMULI EXAMPLE OF EXPANDING TREES FOR BRANCHING AND FOR EACH COMPLEXITY LEVEL, FOR DIGITS AND SYMBOLS. ALSO THE TASK MODIFIED STIMULI IS PRESENTED.

7.1.3 Trial design

Each trial consisted on presenting for 200ms a formula after a fixation point of 500ms. After this the subject would wait between 2800 and 3800 ms for the next trial due to a continuous uniform random jitter of maximum 500ms applied to the onsets of each trial. This jitter is implemented to improve a possible estimation of the HRF (Henson, R. 2007). Moreover the subject could press the button with the right hand at any moment during and after the presentation of the formulas, as requested by the task. In Figure 8 we can appreciate a diagram of the trial design.

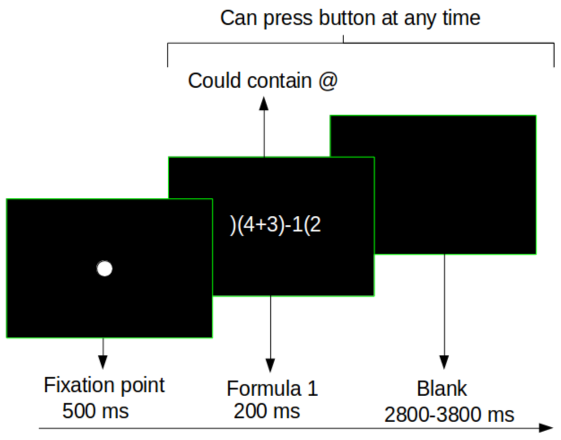


FIGURE 8. TRIAL DESIGN OF EXPANDING TREES

7.2 Behavioral Results

From the analysis of the task accuracy we can see that subjects were able to concentrate on the formulas but had some small difficulties to keep attention for the whole duration of the experiment. Which is reasonable considering the boring nature of the task and the high level of attention required for the quick presentation of the stimuli, as was also reported by the subjects themselves at the end of the experiment. Still in general we can see a mean global accuracy for all subjects of 90.62%, with a a maximum of 100% for some subjects and only two extreme low cases down to 20%. In Figure D.1 in Appendix D we can appreciate the global and individual accuracy of the task

The mean rate of false positives was just of 1.32% with a minimum of 0% for some subjects and a maximum of 10% for few extreme cases. The low false positive rate appreciated for all subjects suggest that there is no difficulty differentiating the modified formulas for the task with respect to normal formulas and that most difficulties at detection would come from attention issues. In Figure D.2 in Appendix D we can appreciate the global and individual false positive rate of the task.

Subjects showing bad extreme values of accuracy or false positive rate should be addressed in an individual analysis of their scans and sanity checks, which has not been done at the moment of writing this report for time constraints.

7.3 FMRI Results

7.3.1 Questions of interest and main contrasts to look at.

In this experimental design we want to contribute to the set of observed activations reported by Maruyama et al. (2012), in particular to see if by simplifying the task we can appreciate increasing complexity effects. We will define the amount of valid operations in the mathematical expression as its complexity level in accordance to the nature of the stimuli designed. Also we will define a mathematical structure to have left or right branching depending on the position of the most inner parentheses (position of longest branch in a tree representation). Moreover we will look at the network of brain areas activated for mathematics, we will look into complexity effects not reported by Maruyama and will also explore the difference in activation between formulas with symbols and digits.

Questions list:

1. Sanity check on the acquisition with the button press.
2. Which brain areas respond to well structured mathematical formulas?
3. Are there areas that respond to unstructured stimuli more than to structured formulas?
4. Do we observe decreasing complexity effects like the ones shown by Maruyama?

5. Do we observe increasing complexity effects not found by Maruyama?
6. Do we see effects of branching in visual areas as Maruyama?
7. Are there areas that respond differently to symbols and digits in formulas?
8. Is there an interaction between complexity and the character category (digits vs symbols)

Corresponding main contrasts lists:

1. Task modified stimuli with button press over non task stimuli without button press.
2. Level 3 formulas over level 0 formulas.
3. Level 0 formulas over level 3 formulas.
4. Inverse linear contrast from level 0 to level 3.
5. Linear contrast from level 0 to level 3.
6. Right over left branching and vice-versa.
7. Symbols over Digits and vice-versa.
8. Increasing complexity in symbols over increasing complexity in digits and vice-versa.

7.3.2 Group results

1. We can clearly see activations in the sensorimotor cortex. These activations are even significant at FWE p-values. The corresponding activations and statistics can be seen in Figure E.1 in Appendix E
2. We can appreciate left mid temporal areas significant at uncorrected values and with low but not significant FWE p-values at the cluster level. In second place it seems there is some activation in the left supramarginal area. If we look only at symbols we seem similar areas but cant find any effects when looking only at digits. The corresponding activations and statistics for the case of all formulas and for only symbols can be seen in Figure E.2 in Appendix E.
3. We did not observe any significant effect to report.
4. We did not observe any significant effect to report.
5. Considering complexity in all formulas, we can only see a cluster of activation in the left mid temporal areas with a FWE p-value of 0.11. But when we look at complexity for the specific case of digits suddenly we find three interesting clusters with very low FWE p-values, one of them overlapping with a greater area on the previously reported left mid temporal area and two more in the right inferior temporal area and the left frontal superior medial area. The corresponding activations and statistics for the case of all formulas and for only digits can be seen in Figure E.3 in Appendix E.
6. We did not observe any significant effect to report.
7. Even though activations are far from being FWE p-value significant, there seems to be an interesting pattern of activation when comparing symbols and digits. We see that activation for symbols have one prominent cluster of activation peaking at a close location to the one reported for the Visual Word Form Area and that this activation is bilateral. Moreover there is a less prominent cluster of activation in the right Insula. In the case of digits we see the activation of two clusters lateralized in the left mid temporal area. The corresponding activations and statistics for the case of symbols and digits can be seen in Figure E.4 in Appendix E.
8. When comparing complexity in symbol formulas over digit formulas we find a peak of FWE p-value significant activity ($p<0.05$) at the left inferior temporal area. In the case of complexity in digits over symbols we observe nothing to report. The corresponding activations and statistics can be seen in Figure E.5 in Appendix E.

7.4 Discussion

First, The button press contrast confirms the quality of the acquisition, giving us more confidence to evaluate the results. Secondly we will discuss the activations we see on the general mathematics network of brain areas and other important trends in the results.

Regarding the mathematics network that we extract by comparing formulas against non formulas (level 3 vs level 0) we observed that it is left lateralized like the language network. Nonetheless it is interesting to see that the specific areas of activation partially differ. In Figure 9 we can see the overlap of the math network for all formulas and only for symbols contrasted with the language network, projected on the mean normalized anatomy of the 16 subjects. From the image becomes clear (although requiring further confirmation) the deviation of the mathematical network and that this deviation is persistent even when considering only formulas with symbols.

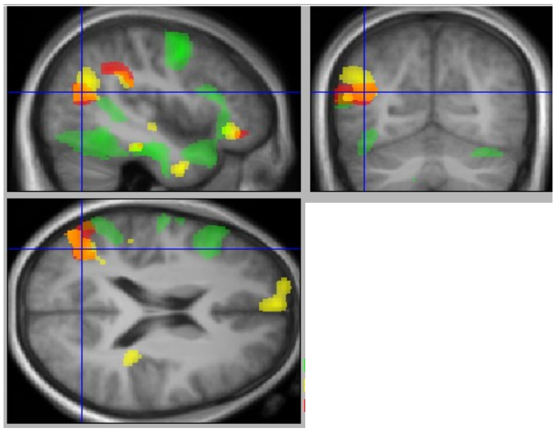


FIGURE 9. OVERLAP OF BRAIN ACTIVATIONS FOR THE MATHEMATICS AND LANGUAGE NETWORK OF ACTIVATED AREAS. THE LANGUAGE AREAS (IN GREEN) AND THE MATHEMATICAL AREAS (IN RED FOR ALL FORMULAS AND IN YELLOW FOR FORMULAS ONLY WITH SYMBOLS). THE INTERSECTION IS SHOWN BY THE MIX OF COLORS. (UNCORRECTED P-VALUE < 0.001)

Also, we can appreciate two important trends in the results. The first trend is related to the fact that we were not able to confirm any of the Maruyama et al. (2012) results for decreasing complexity or branching effects. This might suggest that the difference between the detection task employed in this experimental design and the memory task employed by Maruyama et al. (2012) have an important impact in the brain activations. Moreover we were able to appreciate complexity effects that Maruyama et al. (2012) were not able to detect, which would require further exploration before making any claims. Nonetheless it is interesting to see how the left mid temporal areas activated for complexity in all formulas is then extended and overlapped by complexity in . Which also achieves FWE corrected significance and even shows additional clusters of activation. This makes us wonder why Maruyama et al. (2012) would not be able to find these areas at all. In Figure 10 we can appreciate the overlap of complexity in digit formulas on the complexity of all formulas, projected on the mean normalized anatomy of the 16 subjects.

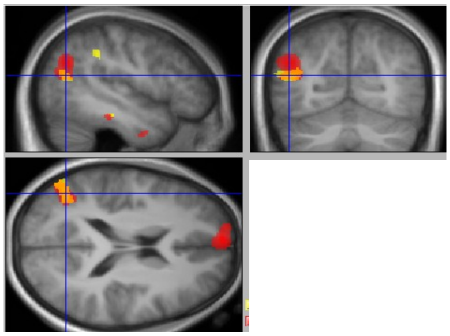


FIGURE 10. OVERLAP OF BRAIN ACTIVATIONS FOR COMPLEXITY IN DIGIT FORMULAS (IN RED) AND COMPLEXITY FOR ALL FORMULAS (IN YELLOW). THE INTERSECTION IS THE MIX OF COLORS (ORANGE). (UNCORRECTED P-VALUE < 0.001)

The second trend is related to the activations we can see when comparing symbols vs digits. In the first place is interesting to see that the activation of symbols appear in coordinates very similar to the reported VWFA (McCandliss, B. D., Cohen, L., & Dehaene, S. 2003) ({-42,-61,-10} in our case vs {-43,-54,-12} reported for the VWFA), alongside a bilateral activation and the activation in the left Insula. In the second place we notice that contrary to symbols, digits show activations lateralized in the left mid temporal area, which is consistent with reports from Shum et al. () on left lateralized areas responsive to digits, nonetheless our areas are a bit off the main reported digit area ({51,-29,-4} the closest in our case vs {51,-54,24} for the digit area). Finally the interaction between complexity and the relation of symbols over digits, give us an interesting and FWE corrected significant result in the left mid temporal area, which is congruent with the left lateralization of brain activations related to symbols and language.

8. EXPERIMENTS IN MATHEMATICS (PRIMING TREES)

The Priming Trees experimental design main idea is to take advantage of the effect of repetition suppression appreciated in the HRF in BOLD-FMRI, as employed by Kouider et al. (2007). Basically when we consider the joint activation of two contiguous stimuli that repeat the same feature, their activation should be less than that of two contiguous stimuli in which the feature is varied, assuming all other features are controlled. This experimental design would allow then to test areas were a syntactic priming effect can be detected. Moreover we consider two levels of complexity in formulas, expecting to find some results converging with Maruyama and the Expanding Trees experimental design.

8.1 Experimental design

8.1.1 Stimuli

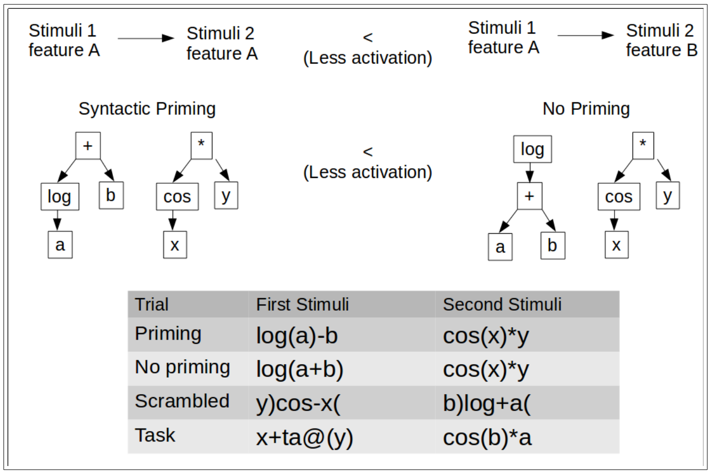
Employing the software Wolfram Mathematica a set of pairs of considered formula patterns was generated to effectively study syntactic priming. Moreover for more robustness we swapped the pairs to have the same population in the roles of prime and target. A detailed example on the generated sets for a pair of formula patterns can be seen in Appendix F in Figure F.1.

In total 10 different formula patterns were considered. 5 of which represent different tree structures, 3 standing for a higher complexity with two mathematical operators and 2 standing for a lower complexity with only one mathematical operator. The other 5 patterns represent scrambled and unstructured formulas that serve as control for the priming that could be caused just due to low level visual differences in the formulas and not due to the change in tree structure or syntactic priming.

All formulas contained uniformly randomly selected symbols from the binary operators {+, -, *}, the unary operators {log, sin, cos, tan} and the pairs of variable symbols {x,y} or {a,b}, never mixing variables from the given groups. Such that all pairs of formulas would contain different operators and variables from each other. Also all formulas contained one pair of parentheses.

By making the formulas different regarding all low level visual characteristics we assume we should effectively be priming only from a syntactic point of view. Only formulas with the same level of complexity and structural category were paired. Such that scrambled was only paired with scrambled for a given complexity level and structured was only paired with structured for a given complexity level. In Appendix F, Table F.1 a complete summary of all formula patterns is given. Moreover in Table F.2 a complete summary with examples of the presented pairs of formulas and their proportion of presentation in each run of acquisition is given. In Figure 11 we present an example of the application of the concept of priming to tree structure and also of scrambled and task modified stimuli.

As a last remark, the presented formulas in the screen comprised 0.72° of vertical visual angle and maximum 3.05° of horizontal visual angle, considering that the largest formula contains 8 characters.



Trial	First Stimuli	Second Stimuli
Priming	log(a)-b	cos(x)*y
No priming	log(a+b)	cos(x)*y
Scrambled	y)cos-x(	b)log+a(
Task	x+ta@(y)	cos(b)*a

FIGURE 11. STIMULI EXAMPLE OF PRIMING TREES. PRIMING EXPLANATORY DIAGRAM AT THE TOP AND STIMULI EXAMPLES FOR PRIMING, NO PRIMING, SCRAMBLED AND TASK CASES AT THE BOTTOM.

8.1.2 Task

As explained in the previous shared methods section. The task is simply to detect an "@" symbol inside the presented mathematical stimulus. In which case, the subject had to press a button with the right hand as soon as possible.

8.1.3 Trial design

Each trial consisted on presenting a fixation point during 500ms, after which two formulas would be presented during 200ms each, separated by a blank screen presented for 100ms. At the end, the subject would wait between 2500 and 3500 ms for the next trial due to a continuous uniform random jitter of maximum 500ms applied to the onsets of each trial. This jitter is implemented to improve a possible estimation of the HRF (Henson, R. 2007). Moreover the subject could press the button with the right hand at any moment during and after the presentation of the formulas, as requested by the task. There were 4 runs of acquisition and in each of them 90 trials were presented. In Figure 12 we can appreciate a diagram of the trial design.

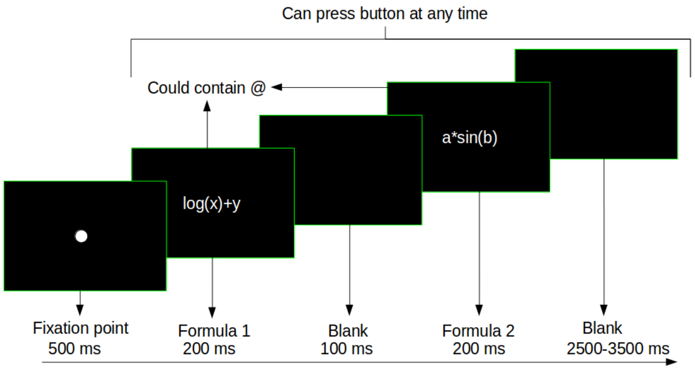


FIGURE 12. TRIAL DESIGN OF PRIMING TREES

8.2 Behavioral Results

As in the case of the Expanding Trees experimental design, from the analysis of the task accuracy we can see that subjects were able to concentrate on the formulas but had some small difficulties to keep attention for the whole duration of the experiment. We can see a mean global accuracy for all subjects of 93.43%, with a a maximum of 100% for some subjects and a minimum of 60%. In Figure G.1 we can appreciate the accuracy of the task, in appendix G.

The mean rate of false positives was just of 0.93% with a minimum of 0% for some subjects and a maximum of 3.75%. The low false positive rate appreciated for all subjects suggest that there is no difficulty differentiating the modified formulas for the task with respect to normal formulas and that any difficulties at detection would come from attention issues. In Figure G.2 we can appreciate the false positive rate of the task, in appendix G.

It is interesting to notice that subjects performed better at this experimental design that at the Expanding Trees design in terms of both accuracy and false positive rate. This could be due to the fact that this experimental design is presented afterwards, so the subjects had time to practice the task, also because the formulas in themselves are easier to process since they have smaller visual angles of presentation and only valid structure.

8.3 Fmri Results

8.3.1 Questions of interest and main contrasts to look at.

In this experimental design we will look into possible convergent evidence on the complexity effects reported in the previous experimental design Expanding Trees. Moreover we will look at the network of brain areas activated for mathematics and into syntactic effects.

Questions list:

1. Sanity check on the acquisition with the button press.
2. Which brain areas respond to well structured mathematical formulas?
3. Are there areas that respond to unstructured stimuli more than to structured formulas?
4. Can we see any decreasing complexity effects?
5. Can we find some increasing complexity effect as shown in Expanding Trees?
6. Can we find brain areas that respond to the syntactic priming effect?
7. What can we find related to the interaction between syntactic priming and structure in formulas?
8. Can we find brain areas responding to specific syntactic structures in the pair of presented formulas?

Corresponding contrasts lists:

1. Task modified stimuli with button press over non task stimuli without button press.
2. Structured formulas over scrambled formulas.
3. Scrambled formulas over structured formulas.
4. Higher complexity formulas (2 operators) over lower complexity formulas (1 operator).
5. Lower complexity formulas (1 operator) over higher complexity formulas (2 operators).
6. Syntactically different formulas over syntactically similar formulas.
7. Interaction between Syntactically different formulas over syntactically similar formulas and structured over non structured formulas.
8. Comparison across pairs of 2-operators-structured formulas with the same syntax. For example a pair of left branched formulas over a pair of right branched formulas. We have three syntactic categories: left branch, right branch and top, that depend on the position of the unary operator in the theoretically parsed tree.

8.3.2 Group results

1. We can clearly see activations in the sensorimotor cortex. These activations are even significant at FWE p-values. The corresponding activations and statistics can be seen in Figure H.1 in Appendix H.
2. We can not observe brain activations related to structured formulas over non structured formulas. This might be due to subject variability or lack of processing of the formulas as a mathematical concept.
3. We only see one non significant possible cluster of activation in the right frontal superior medial area. Along the results of the previous contrast, this might just be noise, so it will be discarded from further analysis on this report.
4. In the case of increasing complexity we only appreciate one cluster of activation in the cingulum anterior area with a low cluster FWE p-value (0.054). The corresponding activations and statistics can be seen in Figure H.2 in Appendix H.
5. In the case of decreasing complexity we see two possible clusters of activation although with poor FWE corrected significance. One in the left angular area and the second in the right temporal inferior area, near the digit area reported by Shum et al. (2013). The corresponding activations and statistics can be seen in Figure H.2 in Appendix H.
6. We can observe two possible clusters of activation, one in the left precentral area and the other in the right mid temporal area, considering their activation peaks low FWE p-values (0.16 and 0.22 respectively). The corresponding activations and statistics can be seen in Figure H.3 in Appendix H.
7. The two possible clusters of activation observed are too small and have too high FWE p-values to be interpretable beyond noise, moreover they are located unexpectedly, one not seen in the rest of the analysis (the left lingual) and the other not even in the cortex (the left thalamus). So we will consider this contrast as not deserving further analysis. The corresponding activations and statistics can be seen in Figure H.3 in Appendix H.
8. Only two contrasts showed activations worth looking into. The first relates left over right positioning of the unary operator in a pair of structured formulas. Two clusters are revealed, one in the right frontal superior medial area and the other in the right mid temporal area, considering their low FWE p-values (0.096 and 0.29 respectively). The second explores the effect of having a pair of formulas where the syntactic structure changes from the unary operator being on one branch to being on the top of the theoretically parsed tree, with respect to pairs of formulas in which there was no syntactic change. We see two interesting clusters, one in the right mid temporal area and the other in the left anterior cingulum , although their FWE p-values are high (0.39 and 0.41 respectively). The corresponding activations and statistics for both menationed contrasts can be seen in Figure H.4 in Appendix H.

8.4 Discussion

First, The button press contrast confirms the quality of the acquisition, giving us more confidence to evaluate the results. Secondly we will comment on the lack of activations seen for a network of brain areas responding to the mathematical stimuli and other important trends in the results.

It caught our attention importantly the fact that we could not differentiate statistically the processing of the pairs of formulas with structure and without structure. This might be do to the lack of additional preprocessing on the data based on individual analysis and unidentified problems with some sessions on some subjects if the effect size of the difference is very small. This issue requires further exploration and analysis.

Moreover, regarding complexity, formulas with bigger trees over formulas with smaller trees, we see in this experimental design areas that do not overlap with those that we report on the previous analyzed experimental design Expanding Trees. In Figure 13 we can appreciate the lack of overlap. It is interesting that both experimental design reveal posterior areas but they are concentrated in different contiguous areas (the cingulum in the case of the Priming Trees experimental design and the left frontal superior medial area in the case of Expanding Trees). In the case of decreasing complexity, it is interesting that we find a peak of activation in the coordinates {51, -44, -18} that is close to the digit area reported by Shum et al. (2013) in {51,-54,24}.

The different areas identified in the case of syntactic priming and syntactic effects related to the processing of pairs of formulas with different and similar structures presented in the previous section are not consistent and not significant enough to allow any particular interpretation at the moment of elaborating this report. It will necessary to dig further into individual analysis of the subjects. The lack of consistency could also be due to just random noise, since the structures might not be processed at all due to the nature of the task. Which is also suggested by the lack of activations in the contrast of structured formulas over non structured ones and then the expected lack of activations in the interaction between syntactic and structural manipulations of the stimuli. Furthermore we have to emphasize that the influence of the presence of parentheses in the formulas was an important low level factor not controlled in the experiment and that even if we found syntactic effects, these would require further exploration controlling for the parentheses effects. Which was beyond the scope of this research project.

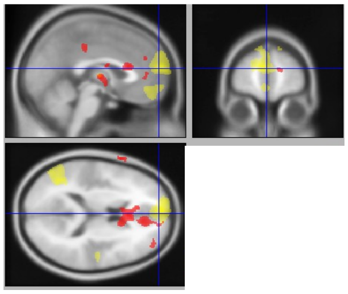


Figure 13. Overlap of brain activations for complexity in the experimental design Priming Trees (in red) and the experimental design Expanding Trees (in yellow). The intersection is the mix of colors (orange). (uncorrected p-value < 0.005)

9. EXPERIMENTS IN LANGUAGE (Analysis of Simple Composition experimental design)

Part of this research project was to analyze the FMRI data acquired for an experimental design developed before my arrival to the laboratory. The idea of this experiment was to create a more sophisticated version of the design implemented by Bemis et al. (2011) to identify brain activation effects related to the composition of phrases in natural language. It was relevant then for me to analyze this experiment to compare some of its results in complexity in natural language with those that I would obtain with my experimental designs and the evidence from other experiments like Maruyama et al. (2012)

9.1 Summary of subjects, experimental design and FMRI acquisition

16 subjects were considered in the preliminary analysis of this dataset, without accounting for the individual analysis due to time constraints at the moment of reporting on this research project.

The stimuli of this experimental design can be summarized as phrases and lists comprising 4 text units, with 1 to 4 words and the remaining as non words. The combination of the words in the case of a phrase convey some joint meaning while in the case of list should not be possible make a composition out of the presented words. An example of phrases and lists that can be formed with 8 words for different levels of composition is shown in Figure 14. The task of this experimental design was simply to indicate if a given word was present in the previous presented list or phrase, which is known as a probe task.

Sequence Type	Phrase 1				Phrase 2			
Combination	sous	mon	bateau	rouge	jean	doit	manger	vite
	twgc	mon	bateau	rouge	twgc	doit	manger	vite
	twgc	thg	bateau	rouge	twgc	tyqwc	manger	vite
	twgc	thg	hrtcvb	rouge	twgc	tyqwc	hrtcvb	vite
List	jean	mon	manger	rouge	sous	doit	bateau	vite
	twgc	mon	manger	rouge	twgc	doit	bateau	vite
	twgc	thg	manger	rouge	twgc	tyqwc	bateau	vite
	twgc	thg	hrtcvb	rouge	twgc	tyqwc	hrtcvb	vite

Words

Figure 14. Stimuli example for phrase and list for all possible levels based on 8 different words.

A trial in this experimental design consists on the sequential presentation of the 4 text units that comprise a phrase or list with a given level of composition. Each text unit is presented for 300ms followed by a blank screen of 300 ms. After the last text unit and blank screen, the screen remains blank for an extra 2 s. Finally the probe task is presented for 3.6 s. A diagram of the trial design is shown in Figure 15.

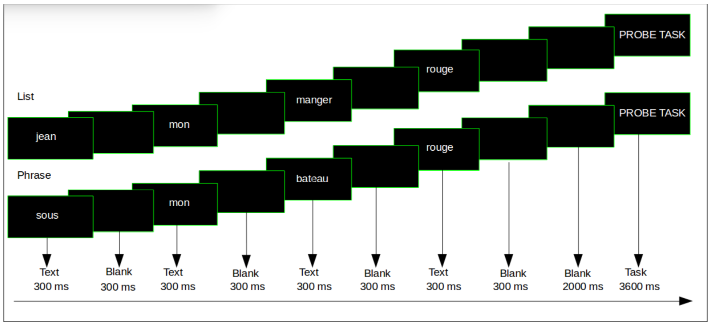


Figure 15. Trial design of Simple Composition experimental design.

9.2 FMRI Results

9.2.1 Questions of interest and main contrasts to look at.

In this experimental design we will look into possible convergent evidence on the complexity effects reported in the previous experimental design Expanding Trees. Moreover we will look at the network of brain areas activated for mathematics and into syntactic effects.

Questions list:

1. What is the approximated network of brain areas processing language?
2. Can we see any complexity effects in the composition of phrases?
3. Can we see a syntactic effect by comparing complexity in phrases over complexity in lists?

Corresponding contrasts lists:

1. Complete phrases over one word sequences.
2. Increasing linear complexity on phrases, from 1 word to 4 words. We can also look at decreasing complexity.
3. Linear complexity of phrases over linear complexity of lists.

9.2.2 Group results

1. We can see activations in the inferior frontal gyrus, left precentral and left mid temporal areas. Which are well known areas for language processing. The corresponding activations and statistics can be seen in Figure I.1 in Appendix I.
2. As in the previous contrast we can mainly see activation in the well known language processing areas. The corresponding activations and statistics can be seen in Figure I.2 in Appendix I.
3. We are not able to see relevant areas of activation. There is only an isolated cluster near the cingulate gyrus. But it just seems like an artifact due to noise. Although we will not further consider this cluster in the analysis, the corresponding activations and statistics are presented in Figure I.3 in Appendix I.

9.3 Discussion

As we would expect, contrasting complete phrases against a sequences of non words ending in one word reveals plenty of the areas normally observed for language as summarized by Friederici (2011). We see the activation of Broca's area in the inferior frontal gyrus (IFG), Wernicke's area in the superior temporal gyrus (STG), as well as parts of the middle temporal gyrus (MTG). In Figure 16 we can see an overlap of our contrast with the language processing areas identified with the language localizer. In the case of increasing phrase complexity we observe even a stronger overlap of the same langauge areas as can be seen in Figure 17. Nonetheless it is interesting to see that phrase complexity emphasizes posterior areas in the network, particularly the left inferior frontal gyrus and the left precentral areas.

We would emphasize that the fact that we were not able to see significant composition effects when comparing phrase and list complexity could be due to the missing extra assessment of individual subject scans and the respective extra processing that could be applied afterwards. Moreover additional methods of estimation and extra sanity checks with a language localizer implemented for this experimental design have not been conducted due to time constraints at the moment of writing this report.

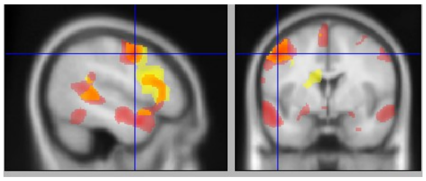


Figure 16. Overlap of brain activations for language as identified in the language localizer (in red) and language as identified in Simple Composition (in yellow). The intersection is the mix of colors (orange). (uncorrected p-value < 0.005)

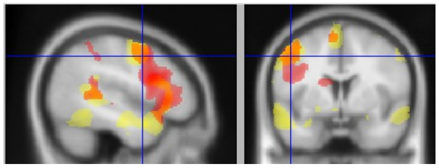


Figure 17. Overlap of brain activations for phrase complexity (in red) and language as identified in the language localizer (in yellow). The intersection is the mix of colors (orange). (uncorrected p-value < 0.005)

10. CONCLUSIONS AND FURTHER RESEARCH

In summary we were able to appreciate brain areas active in the language network and the sensorimotor cortex due to the language localizer and the button press task that serve as preliminary sanity check for the group analysis. In the case of our first experimental design, Expanding Trees, we were not able to see the same activations reported by Maruyama et al (2012), but could appreciate activations related to increasing complexity. Moreover we could observe differences in the networks of brain areas involved in the processing of formulas with digits and symbols. Finally in the case of the second experimental design, Priming Trees, we were not able to observe similar activations to the first experimental design and we could not even observe a difference between structured and unstructured mathematical expressions.

Nonetheless as mentioned in a previous occasion there are still 4 missing subjects, scheduled to be scanned soon, in our analysis. Moreover we still need to realize an individual analysis and extra sanity checks on the individual data to assess the need to discard subjects, runs of acquisition or apply extra preprocessing steps to the collected data. Furthermore we should explore other possible preprocessing pipelines as proposed by Strother (2006) and possibly other estimation methods for the HRF and the detection of its activations. We should also improve the group analysis with more sophisticated models that could asses better the strong inter-subject variability present in BOLD-FMRI studies.
Considering further research, in the case of the complexity of

tree structures we would like to test complexity in number of constituents without the repairing effects that might be taking place under the Maruyama's stimuli paradigm due to the ungrammatical portions in lower level formulas. Moreover we would like to find a way to test for the composition of Mathematical constituents and link this with a complexity effect. We also would like to refine the experimental designs to be able to account for the grouping effects introduced by parentheses.

Finally we have some ideas on implementing in neuroimaging studies in a similar line to the spatial and phrasal effects studied by Goldstone and Jansen respectively. Going even further, we have the idea of employing Paul Smolensky's (2006) framework of connectionist computation to model mathematical symbolic representations, complexity and composition. So we can test modeled predictions of the inner workings of the neural circuits responsible for the parsing of mathematical formulas and calculation with new neuroimaging experiments and experimental paradigms still not explored in the field.

11. ACKNOWLEDGEMENT

Thanks to Dr Christophe Pallier for his dedication and commitment as supervisor of this project, encouraging and discouraging me whenever needed to keep me on the right track. Thanks to Dr. Stanislas Dehaene for giving me the initial opportunity to work at the Unicog lab in NeuroSpin. Thanks to Dr Khashayar Pakdaman for his dedication as director of the master's program and assistance in the successful development of all research projects in the master, including this one. Finally thanks to Esther Lin and Murielle Fabre members of the team for their help and to all other fellow students, researchers and professionals at NeuroSpin that warmly coped with all my doubts, research efforts and endless, sometimes annoying, curiosity.

12. REFERENCES

[1] Baker, M. *The Atoms of Language*, Basic Books, 2001.

[2] Bemis, Douglas K., and Liina Pylkkänen. "Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases." *The Journal of Neuroscience* 31.8 (2011): 2801-2814.

[3] Ben-Shachar M[1], Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y.(2003) "The neural reality of syntactic transformations evidence from functional magnetic resonance imaging." *Psychological Science* 14(5) 433-440.

[4] Ben-Shachar, Michal, Dafna Palti, and Yosef Grodzinsky. (2004) "Neural correlates of syntactic movement: converging evidence from two fMRI experiments." *Neuroimage* 21(4): 1320-1336.

[5] Börgers C and Kopell N (2003). "Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity." *Neural computation* 15(3) : 509-538.

[6] Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. The journal of neuroscience, 16(13), 4207-4221.

[7] Branigan H (2007). "Syntactic priming." *Language and Linguistics Compass* 1 (1-2): 1–16

[8] Buckner, R.L., 1998. Event-related fMRI and the hemodynamic response. Hum. Brain Mapp. 6 (5–6), 373 – 377.

[9] Chaâri, L., Menauz, S., Pesquet, J., & Ciuciu, P. (2010, November). Impact of the parallel imaging reconstruction algorithm on brain activity detection in fMRI. In Applied Sciences in Biomedical and Communication Technologies

[10] Chen, W., & Ogawa, S. (1996). 10 Principles of BOLD Functional MRI. Red, 10, 1.

[11] Chomsky, N. *Syntactic Structures,* La Haye: Mouton, 2000

[12] David J.T, and Bever TG (2001). *Sentence comprehension: The integration of habits and rules*. Vol. 1950. Cambridge, MA: MIT Press

[13] Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. Trends in cognitive sciences, 15(6), 254-262.

[14] Devauchelle AD, Oppenheim C, Rizzi L, Dehaene S, Pallier C. (2009) "Sentence syntax and content in the human temporal lobe: an fMRI adaptation study in auditory and visual modalities." *Journal of cognitive neuroscience* 2(.5) 1000-1012.

[15] Frank SL and Bod R (2011) "Insensitivity of the human sentence-processing system to hierarchical structure." *Psychological Science* 22(6): 829-834.

[16] Friederici, A. D. (2011). The brain basis of language processing: from structure to function. Physiological reviews, 91(4), 1357-1392.

[17] Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2458-2463.

[18] Friedrich R, Friederici AD (2009): "Mathematical logic in the human brain: Syntax" *Plos One* 4(5) :e5599.

[19] Friedrich, R. and Friederici, A.D.. (2009). Mathematical Logic in the Human Brain: Syntax. *PLoS One,* 4 (5): e5599.

[20] Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage, 21(4), 1639-1651.

[21] Henson, R. (2007). Efficient experimental design for fMRI. Statistical parametric mapping: The analysis of functional brain images, 193-210.

[22] Izhikevich EM. "Polychronization (2006): computation with spikes." *Neural computation* 18.2:: 245-282.

[23] Jackendorf, R. *Patterns in the mind: Language and human nature*. Basic Books, 2008.

[24] Jansen R, Marriott K, and Yelland GW (2007) "Parsing of algebraic expressions by experienced users of mathematics." *European Journal of Cognitive Psychology* 19(2): 286-320.

[25] Jansen R., Marriott K, and Yelland GW (2003). "Comprehension of algebraic expressions by experienced

users of mathematics." *The Quarterly Journal of Experimental Psychology:* 56(1) 3-30.

[26] Jansen, A. R., Marriott, K., & Yelland, G. W. (2003). Comprehension of algebraic expressions by experienced users of mathematics. The Quarterly Journal of Experimental Psychology: Section A, 56(1), 3-30.

[27] Jansen, A. R., Marriott, K., & Yelland, G. W. (2007). Parsing of algebraic expressions by experienced users of mathematics. European Journal of Cognitive Psychology, 19(2), 286-320.

[28] Jansen, A.R., Marriott, K. & Yelland, G.W. (2000). Constituent Structure in Mathematical Expressions.â In Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society 13(15) 22:238. Psychology Press.

[29] Jurafsky D, and Martin JH. *Speech & Language Processing*. Pearson Education India, 2000.

[30] Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. Trends in cognitive sciences, 6(8), 350-356. (ISABEL), 2010 3rd International Symposium on (pp. 1-5). IEEE.

[31] Kouider, S., Dehaene, S., Jobert, A., & Le Bihan, D. (2007). Cerebral bases of subliminal and supraliminal priming during reading. Cerebral Cortex, 17(9), 2019-2029.

[32] Krause, F and Lindemann, O. (2014). Expyriment: A Python library for cognitive and neuroscientific experiments. Behavior Research Methods, 46(2), 416-428.

[33] Landy, D., & Goldstone, R. L. (2007a). Formal notations are diagrams: Evidence from a production task. Memory & Cognition, 35(8), 2033-2040.

[34] Landy, D., & Goldstone, R. L. (2007b). How abstract is symbolic thought?. Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(4), 720.

[35] Lindquist, M. A. (2008). The statistical analysis of fMRI data. Statistical Science, 23(4), 439-464.

[36] Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. The Journal of Neuroscience, 23(10), 3963-3971.

[37] Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. Nature, 453(7197), 869-878.

[38] Maruyama M , Pallier C , Jobert A, Sigman M, Dehaene S (2012): "The cortical representation of simple mathematical expressions" NeuroImage 61(4):1444-1460

[39] Maruyama, M., Pallier, C., Jobert, A., Sigman, M. & Dehaene, S. (2012). The Cortical Representation of Simple Mathematical Expressions. Neuroimage 61 (4): 1444-60.

[40] McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. Trends in cognitive sciences, 7(7), 293-299.

[41] Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, Büchel C, Weiller C (2003) "Broca's area and the language instinct." *Nature neuroscience* 6(7): 774-781.

[42] Nespor M, and Vogel I. *Prosodic phonology: with a new foreword*. Vol. 28. Walter de Gruyter, 2007.

[43] Noppeney U and Price C.J (2004). "An fMRI study of syntactic adaptation." *Journal of Cognitive Neuroscience* 16(4): 702-713.

[44] Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proceedings of the National Academy of Sciences, 89(13), 5951-5955.

[45] Oldfield RC (1971)" the assessment and analysis of handedness: The edinburg inventory" *Neuropsychologi*a, 9: 97-113

[46] Pallier C, Devauchelle A.-D and Dehaene S. (2011). Cortical Representation of the Constituent Structure of Sentences. *Proceedings of the National Academy of Science USA* 108 (6): 2522-27.

[47] Pallier C, Devauchelle AD, Dehaene S ." (2011) Cortical representation of the constituent structure of sentences "*Proceedings of the National Academy of Sciences* 108(6):2522-2527

[48] Plate, Tony A (1995). "Holographic reduced representations." *Neural networks, IEEE transactions on* 6(3): 623-641.

[49] Pollack, Jordan B (1988). "Recursive auto-associative memory." *Neural Networks* 1: 115-122.

[50] Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage, 62(2), 816-847.

[51] Shum J[1], Hermes D, Foster BL, Dastjerdi M, Rangarajan V, Winawer J, Miller KJ, Parvizi J (2013). A brain area for visual numerals, *Journal of neurosciences* 33 (16):6709-15

[52] Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ,Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney D.E, Niazy R, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, and Matthews PM (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23(S1):208-219

[53] Smolensky, P, & Legendre, G. (2006). The Harmonic Mind Vol. 1. MIT Press.

[54] Strother, S. C. (2006). Evaluating fMRI preprocessing pipelines. Engineering in Medicine and Biology Magazine, IEEE, 25(2), 27-41.

[55] Tettamanti M, Alkadhi H, Moro A′ Perani D′ Kollias S, (2002)"Neural correlates for the acquisition of natural language syntax." *Neuroimage* 17(2): 700-709.

[56] Tyler, L. K., & Marslen-Wilson, W. (2008). Fronto-temporal brain systems supporting spoken language comprehension. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1493), 1037-1054.

[57] Wager T. D and Nichols T. E. (2003). Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage*, *18*(2), 293-309.

[58] Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. Human brain mapping, 4(1), 58-73.

Appendix A. Details on the developed pipeline

The automatized pipeline was developed in Bash Scripting and Matlab code to be run on the Linux O.S. The pipeline mainly exploits standard algorithms contained in diverse packages like FSL (Smith et al, 2004), SPM8 (http:// www.fil.ion.ucl.ac.uk/spm/) and PyPreprocess (http://github.com/neurospin/pypreprocess). This common pipeline was developed with the aim of replacing easily Matlab code in the future by python implementations of the processing and preprocessing steps. Also it would allow easy update of the analysis as subjects are scanned and would save time by allowing all experimental designs to be analyzed under a common pipeline.

In the follwing diagrams, more details are given in the conceptualized file structure that determines how files have to be constructed and passed to the pipeline and what files and folders are automatically generated by running the pipeline. The pipeline can be simply run by calling the script automatic_pipeline_update.sh from a linux terminal, as long as the necessary files are given it will run the complete analysis only for those subjects for which it has not be done and recomputes the group analysis. Details of the code are not provided in the appendix since they would comprise a disproportionately sized annex on a topic that is not the main focus of interest of this research project.

A.1) File structure of files necessary to run the pipeline

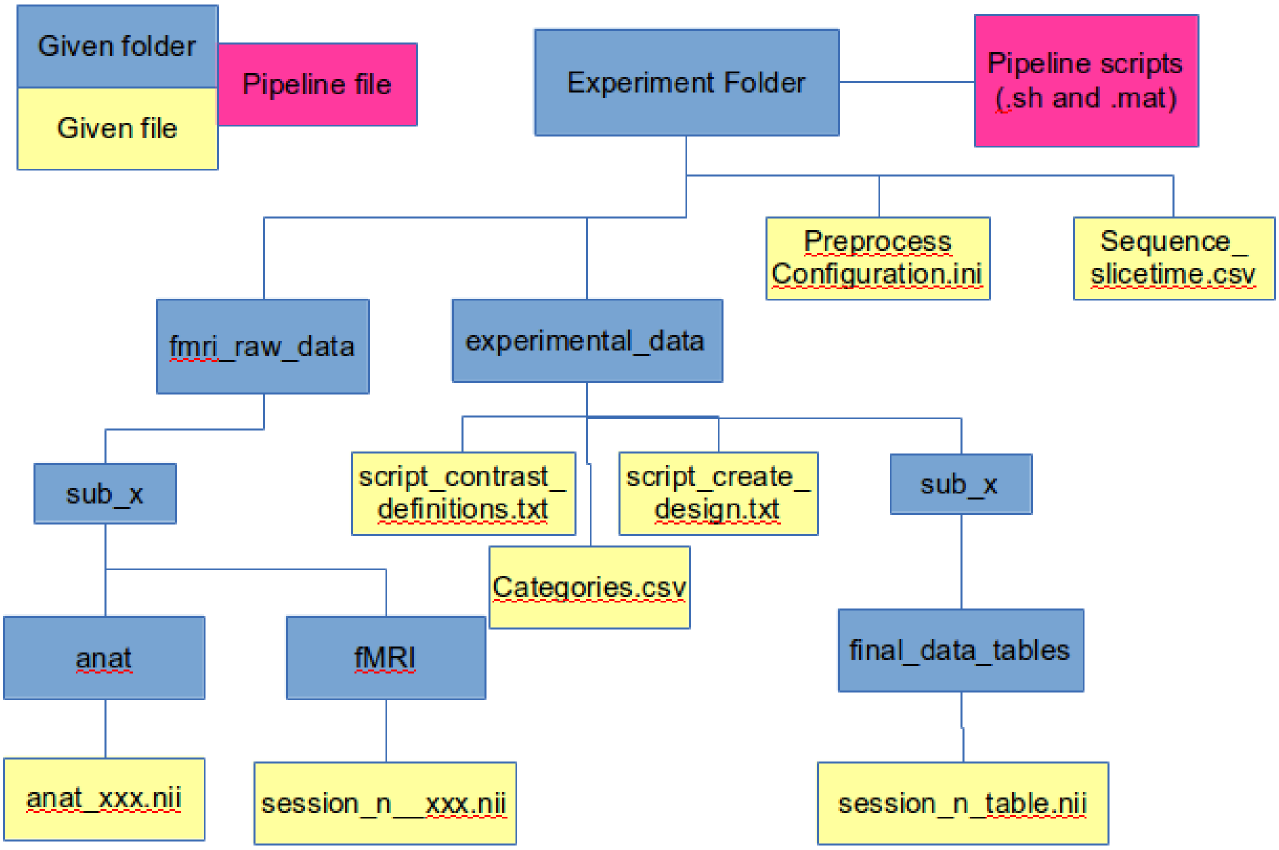


Figure A.1. File structure of files necessary to run the pipeline

A.2) File structure of files generated by the pipeline

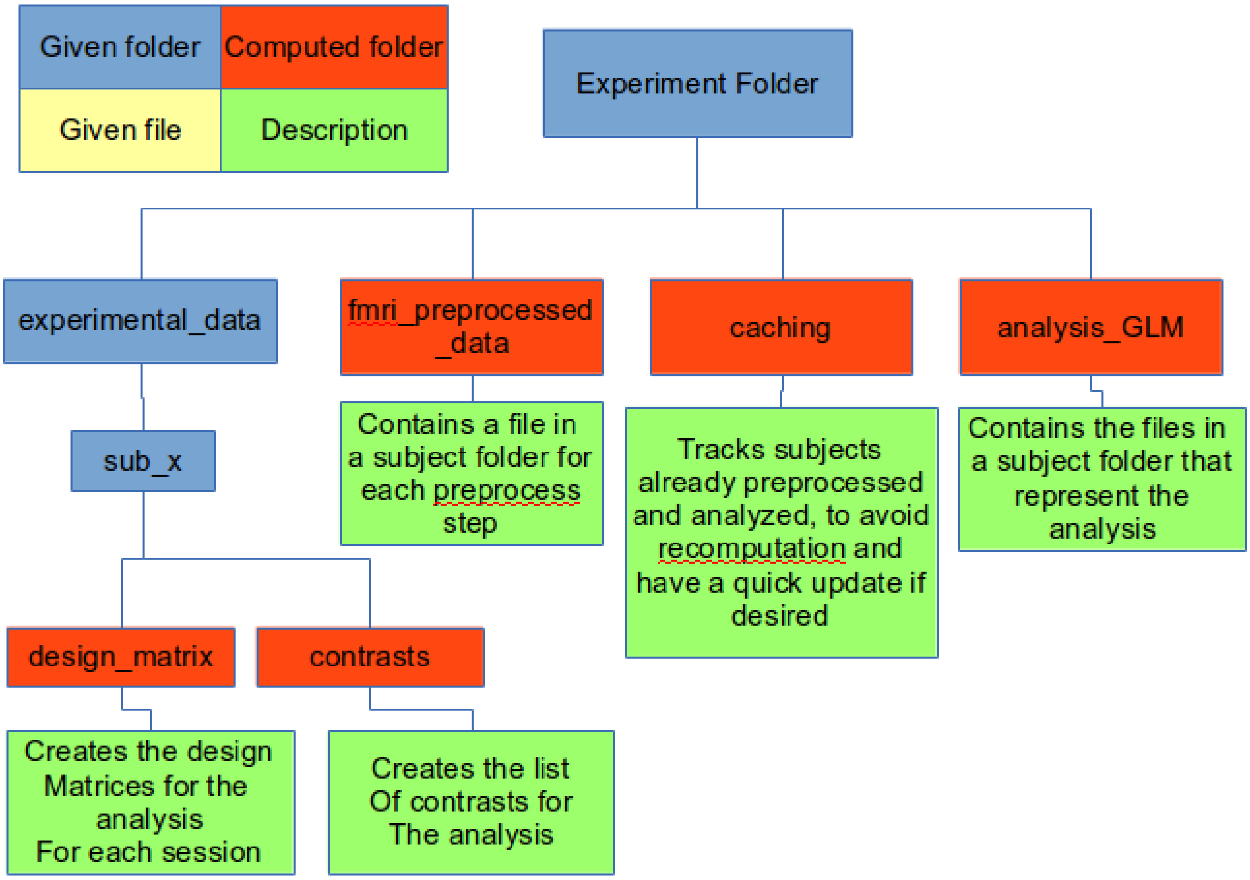
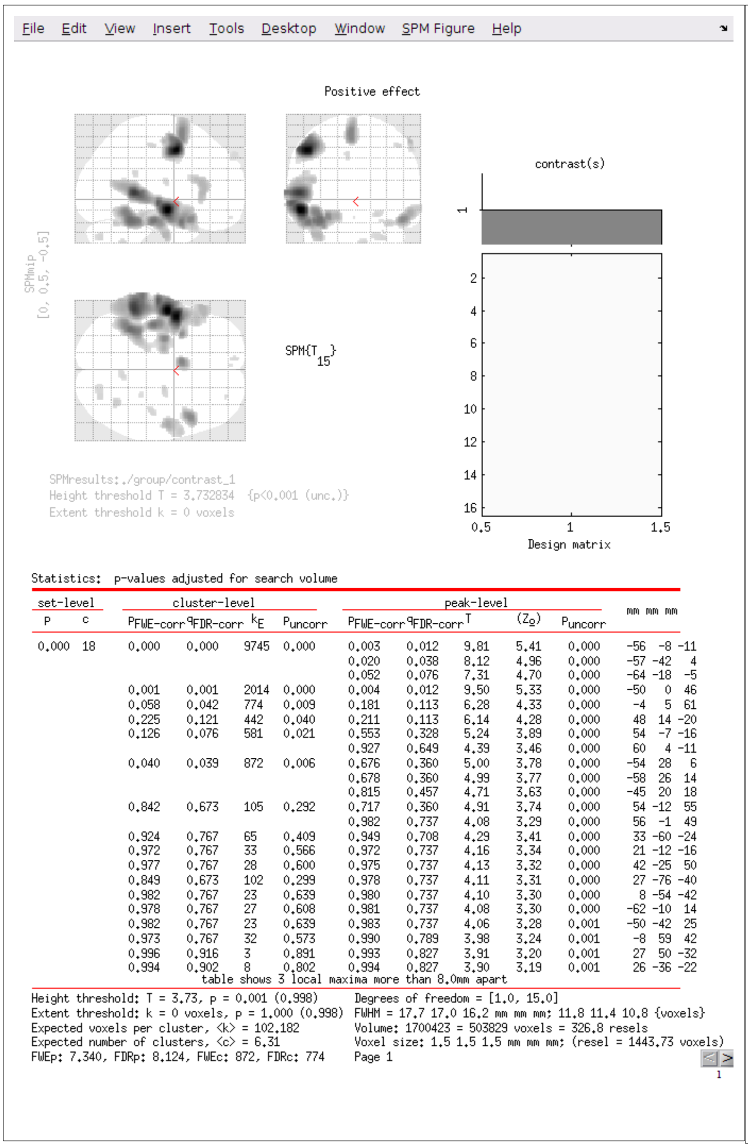


Figure A.2. File structure of files generated by the pipeline

Appendix B. Language Localizer Brain Activations

B.1) Language related areas (p-values < 0.001 to the left and FWE p-values<0.05 to the right)



SPMresults:./group/contrast_1
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

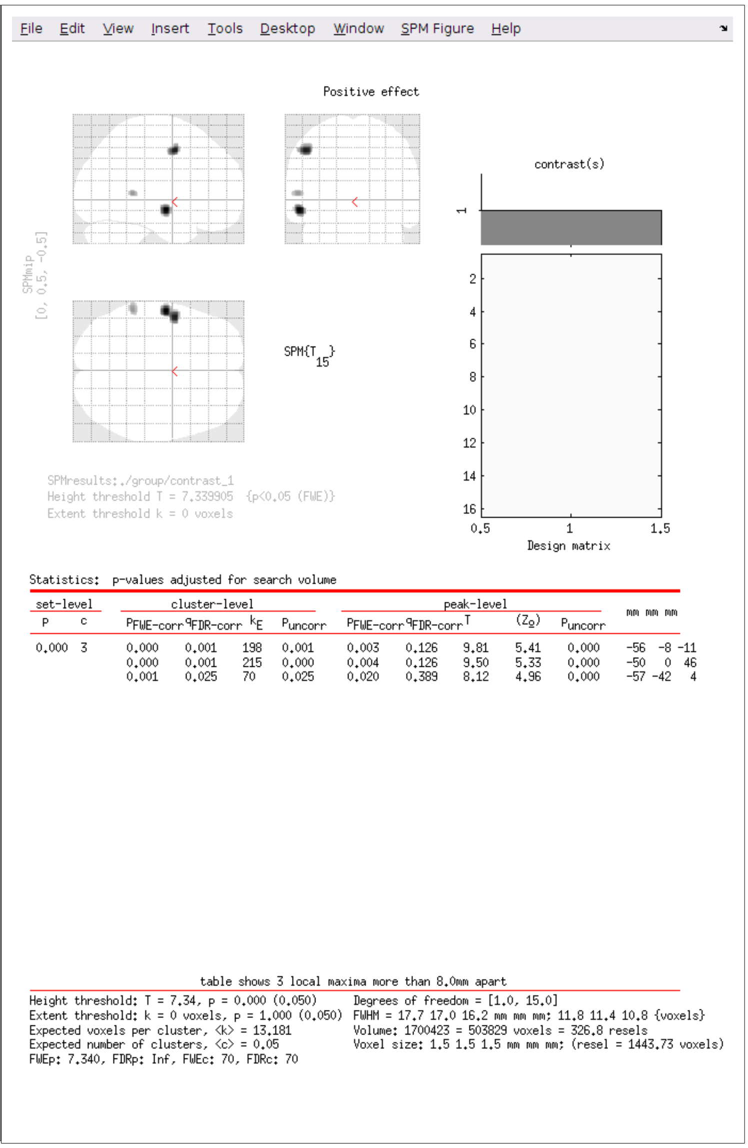
Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level					mm mm mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	
0.000	18	0.000	0.000	9745	0.000	0.003	0.012	9.81	5.41	0.000	-56 -8 -11
						0.020	0.038	8.12	4.96	0.000	-57 -42 4
						0.052	0.076	7.31	4.70	0.000	-64 -18 -5
		0.001	0.001	2014	0.000	0.004	0.012	9.50	5.33	0.000	-50 0 46
		0.058	0.042	774	0.009	0.181	0.113	6.28	4.33	0.000	-4 5 61
		0.225	0.121	442	0.040	0.211	0.113	6.14	4.28	0.000	48 14 -20
		0.126	0.076	581	0.021	0.553	0.328	5.24	3.89	0.000	54 -7 -16
						0.927	0.649	4.39	3.46	0.000	60 4 -11
		0.040	0.039	872	0.006	0.676	0.360	5.00	3.78	0.000	-54 28 6
						0.678	0.360	4.99	3.77	0.000	-58 26 14
						0.815	0.457	4.71	3.63	0.000	-45 20 18
		0.842	0.673	105	0.292	0.717	0.360	4.91	3.74	0.000	54 -12 55
						0.982	0.737	4.08	3.29	0.000	56 -1 49
		0.924	0.767	65	0.409	0.949	0.708	4.29	3.41	0.000	33 -60 -24
		0.972	0.767	33	0.566	0.972	0.737	4.16	3.34	0.000	21 -12 -16
		0.977	0.767	28	0.600	0.975	0.737	4.13	3.32	0.000	42 -25 50
		0.849	0.673	102	0.299	0.978	0.737	4.11	3.31	0.000	27 -76 -40
		0.982	0.767	23	0.639	0.980	0.737	4.10	3.30	0.000	8 -54 -42
		0.978	0.767	27	0.608	0.981	0.737	4.08	3.30	0.000	-62 -10 14
		0.982	0.767	23	0.639	0.983	0.737	4.06	3.28	0.001	-50 -42 25
		0.973	0.767	32	0.573	0.990	0.789	3.98	3.24	0.001	-8 59 42
		0.996	0.916	3	0.891	0.993	0.827	3.91	3.20	0.001	27 50 -32
		0.994	0.902	8	0.802	0.994	0.827	3.90	3.19	0.001	26 -36 -22

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (0.998)
Extent threshold: k = 0 voxels, p = 1.000 (0.998)
Expected voxels per cluster, <k> = 102.182
Expected number of clusters, <c> = 6.31
FWEp: 7.340, FDRp: 8.124, FWEc: 872, FDRc: 774

Degrees of freedom = [1.0, 15.0]
FWHM = 17.7 17.0 16.2 mm mm mm; 11.8 11.4 10.8 {voxels}
Volume: 1700423 = 503829 voxels = 326.8 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1443.73 voxels)
Page 1



SPMresults:./group/contrast_1
Height threshold T = 7.339905 {p<0.05 (FWE)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level					mm mm mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	
0.000	3	0.000	0.001	198	0.001	0.003	0.126	9.81	5.41	0.000	-56 -8 -11
		0.000	0.001	215	0.000	0.004	0.126	9.50	5.33	0.000	-50 0 46
		0.001	0.025	70	0.025	0.020	0.389	8.12	4.96	0.000	-57 -42 4

table shows 3 local maxima more than 8.0mm apart

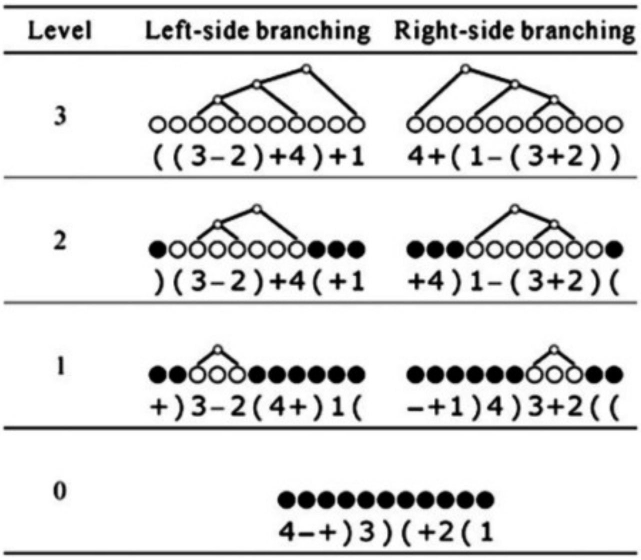
Height threshold: T = 7.34, p = 0.000 (0.050)
Extent threshold: k = 0 voxels, p = 1.000 (0.050)
Expected voxels per cluster, <k> = 13.181
Expected number of clusters, <c> = 0.05
FWEp: 7.340, FDRp: Inf, FWEc: 70, FDRc: 70

Degrees of freedom = [1.0, 15.0]
FWHM = 17.7 17.0 16.2 mm mm mm; 11.8 11.4 10.8 {voxels}
Volume: 1700423 = 503829 voxels = 326.8 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1443.73 voxels)

Figure B.1. Language related areas (p-values < 0.001 to the left and FWE p-values<0.05 to the right)

Appendix C. Expanding Trees Stimuli

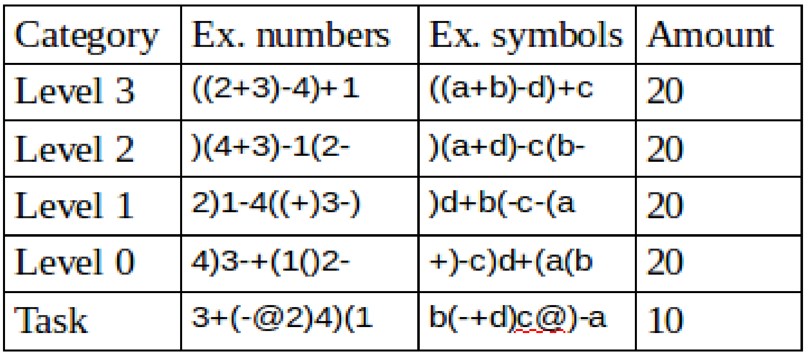
C.1) Stimuli examples for the different levels of complexity and branching, in the case of numbers.



Level	Left-side branching	Right-side branching
3	((3−2)+4)+1	4+(1−(3+2))
2	)(3−2)+4(+1	+4)1−(3+2)(
1	+)3−2(4+)1(	−+1)4)3+2((
0	4−+)3)(+2(1	

Figure C.1. Diagram of stimuli for each level and branching in the case of formulas with numbers. Taken from Maruyama et al.

C.2) Stimuli examples and their proportions



Category	Ex. numbers	Ex. symbols	Amount
Level 3	((2+3)-4)+1	((a+b)-d)+c	20
Level 2	)(4+3)-1(2-	)(a+d)-c(b-	20
Level 1	2)1-4((+)3-)	)d+b(-c-(a	20
Level 0	4)3-+(1()2-	+)-c)d+(a(b	20
Task	3+(-@2)4)(1	b(-+d)c@)-a	10

Table C.1. Table of stimuli example for all levels for numbers and symbols and the task. Including the proportion of presentation in number of stimuli for each run of acquisition.

Appendix D. Expanding Trees Behavioral results

D.1) Task accuracy

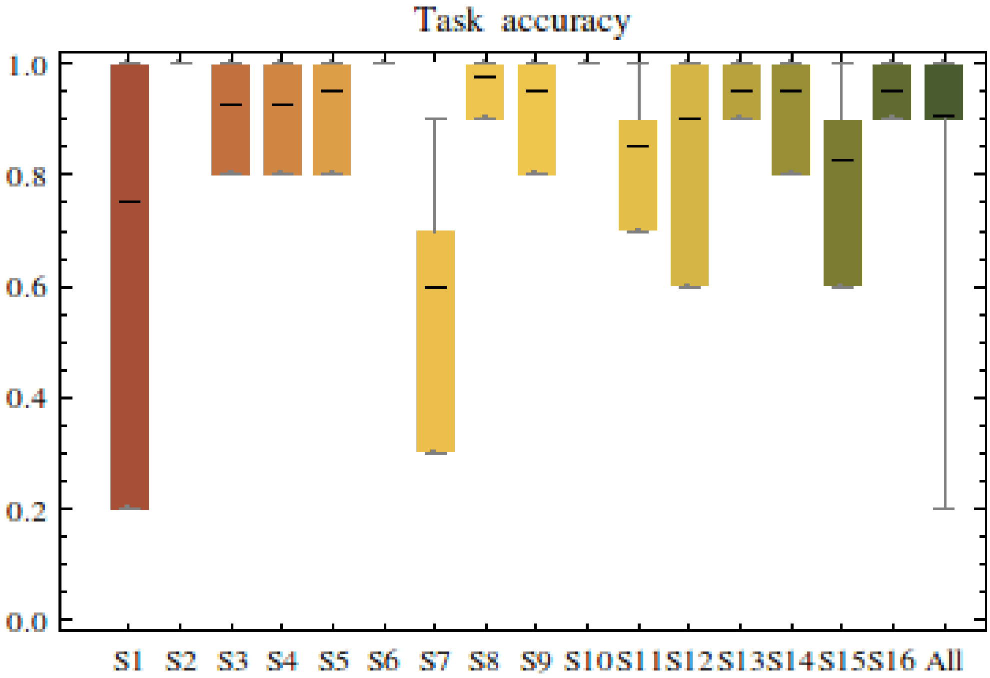


Figure D.1. Accuracy of task in Expanding Trees

D.2) Task false positive rate

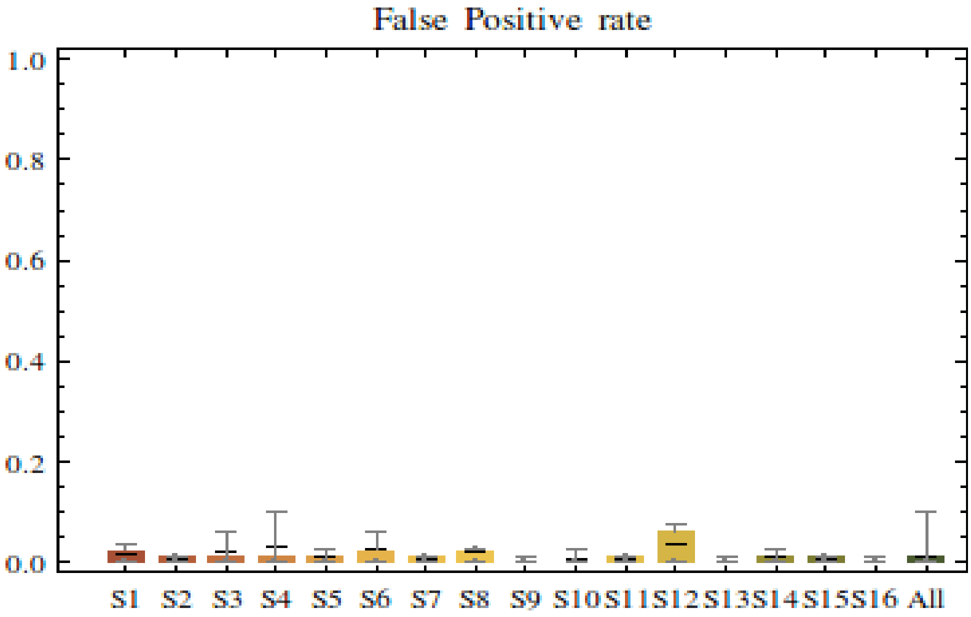
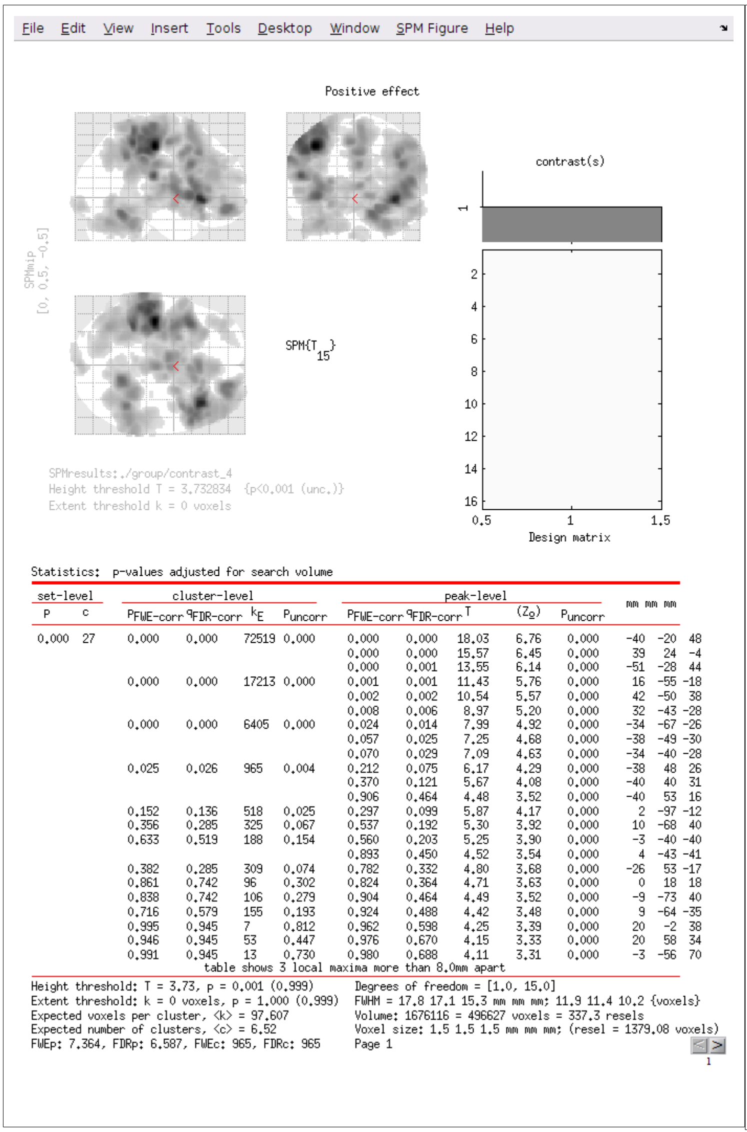


Figure D.2. False positive rate of task in Expanding Trees

Appendix E. Expanding Trees Brain Activations

E.1) Button press: brain glass image and statistic tables for uncorrected p-value < 0.001 and FWE corrected p-value<0.05.



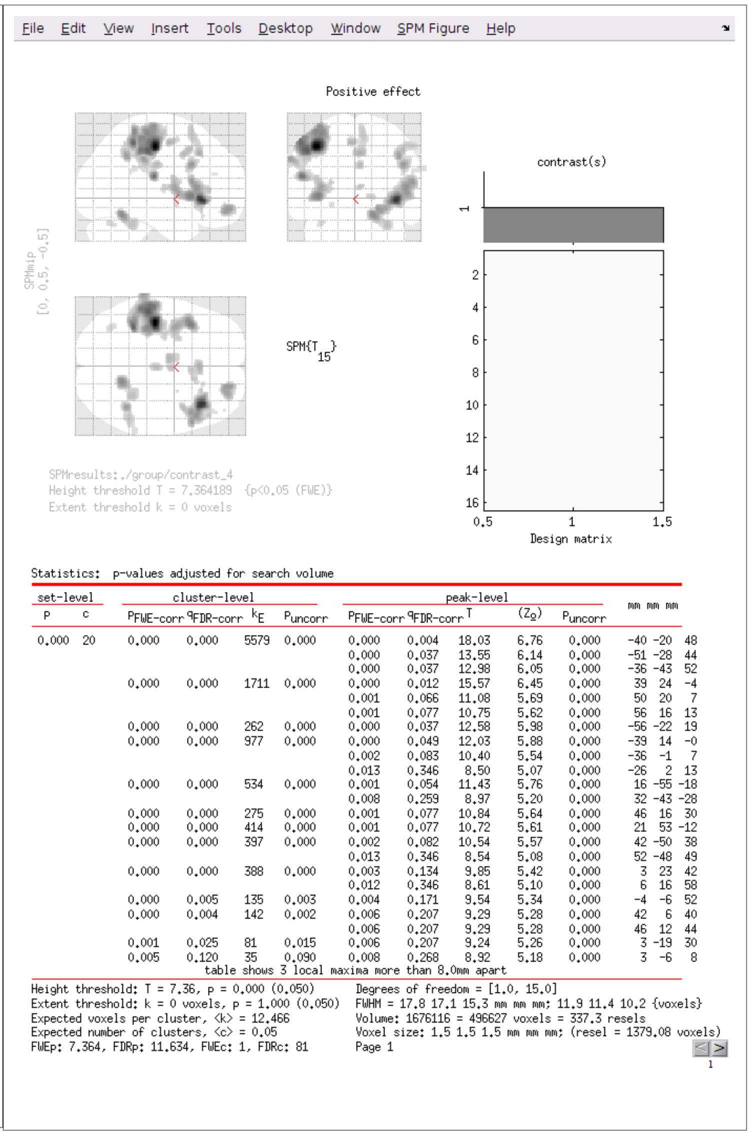
SPMresults: ./group/contrast_4
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	27	0.000	0.000	72519	0.000	0.000	0.000	18.03	6.76	0.000	-40	-20	48
						0.000	0.000	15.57	6.45	0.000	39	24	-4
						0.000	0.001	13.55	6.14	0.000	-51	-28	44
		0.000	0.000	17213	0.000	0.001	0.001	11.43	5.76	0.000	16	-55	-18
						0.002	0.002	10.54	5.57	0.000	42	-50	38
						0.008	0.006	8.97	5.20	0.000	32	-43	-28
		0.000	0.000	6405	0.000	0.024	0.014	7.99	4.92	0.000	-34	-67	-26
						0.057	0.025	7.25	4.68	0.000	-38	-49	-30
						0.070	0.029	7.09	4.63	0.000	-34	-40	-28
		0.025	0.026	965	0.004	0.212	0.075	6.17	4.29	0.000	-38	48	26
						0.370	0.121	5.67	4.08	0.000	-40	40	31
						0.906	0.464	4.48	3.52	0.000	-40	53	16
		0.152	0.136	518	0.025	0.297	0.099	5.87	4.17	0.000	2	-97	-12
		0.356	0.285	325	0.067	0.537	0.192	5.30	3.92	0.000	10	-68	40
		0.633	0.519	188	0.154	0.560	0.203	5.25	3.90	0.000	-3	-40	-40
						0.893	0.450	4.52	3.54	0.000	4	-43	-41
		0.382	0.285	309	0.074	0.782	0.332	4.80	3.68	0.000	-26	53	-17
		0.861	0.742	96	0.302	0.824	0.364	4.71	3.63	0.000	0	18	18
		0.838	0.742	106	0.279	0.904	0.464	4.49	3.52	0.000	-9	-73	40
		0.716	0.579	155	0.193	0.924	0.488	4.42	3.48	0.000	9	-64	-35
		0.995	0.945	7	0.812	0.962	0.598	4.25	3.39	0.000	20	-2	38
		0.946	0.945	53	0.447	0.976	0.670	4.15	3.33	0.000	20	58	34
		0.991	0.945	13	0.730	0.980	0.688	4.11	3.31	0.000	-3	-56	70

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.999)         Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.999)     FWHM = 17.8 17.1 15.3 mm mm mm; 11.9 11.4 10.2 {voxels}
Expected voxels per cluster, <k> = 97.607             Volume: 1676116 = 496627 voxels = 337.3 resels
Expected number of clusters, <c> = 6.52               Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1379.08 voxels)
FWEp: 7.364, FDRp: 6.587, FWEc: 965, FDRc: 965        Page 1
```



SPMresults: ./group/contrast_4
Height threshold T = 7.364189 {p<0.05 (FWE)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

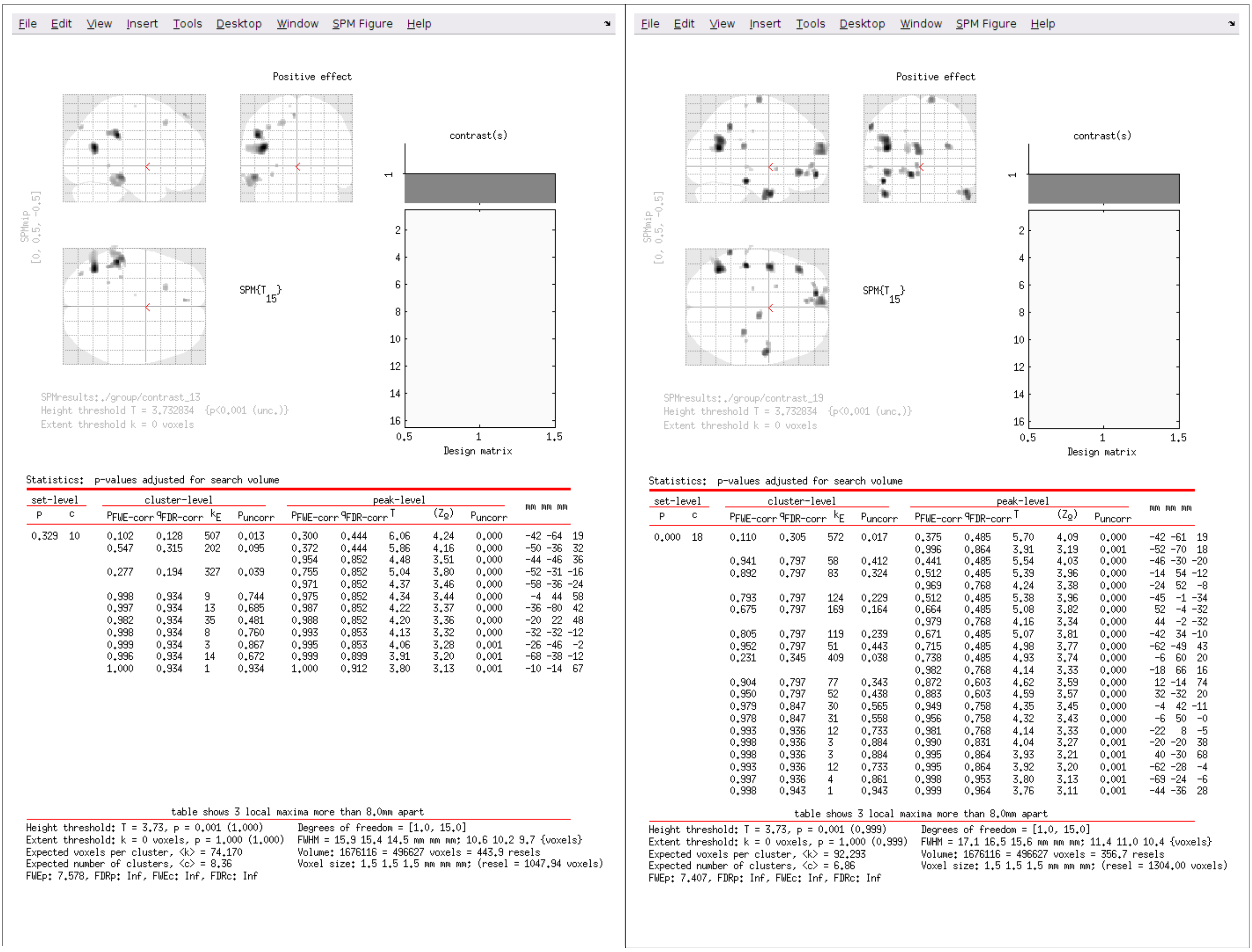
set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	20	0.000	0.000	5579	0.000	0.000	0.004	18.03	6.76	0.000	-40	-20	48
						0.000	0.037	13.55	6.14	0.000	-51	-28	44
						0.000	0.037	12.98	6.05	0.000	-36	-43	52
		0.000	0.000	1711	0.000	0.000	0.012	15.57	6.45	0.000	39	24	-4
						0.001	0.066	11.08	5.69	0.000	50	20	7
						0.001	0.077	10.75	5.62	0.000	56	16	13
		0.000	0.000	262	0.000	0.000	0.037	12.58	5.98	0.000	-56	-22	19
		0.000	0.000	977	0.000	0.000	0.049	12.03	5.88	0.000	-39	14	-0
						0.002	0.083	10.40	5.54	0.000	-36	-1	7
						0.013	0.346	8.50	5.07	0.000	-26	2	13
		0.000	0.000	534	0.000	0.001	0.054	11.43	5.76	0.000	16	-55	-18
						0.008	0.259	8.97	5.20	0.000	32	-43	-28
		0.000	0.000	275	0.000	0.001	0.077	10.84	5.64	0.000	46	16	30
		0.000	0.000	414	0.000	0.001	0.077	10.72	5.61	0.000	21	53	-12
		0.000	0.000	397	0.000	0.002	0.082	10.54	5.57	0.000	42	-50	38
						0.013	0.346	8.54	5.08	0.000	52	-48	49
		0.000	0.000	388	0.000	0.003	0.134	9.85	5.42	0.000	3	23	42
						0.012	0.346	8.61	5.10	0.000	6	16	58
		0.000	0.005	135	0.003	0.004	0.171	9.54	5.34	0.000	-4	-6	52
		0.000	0.004	142	0.002	0.006	0.207	9.29	5.28	0.000	42	6	40
						0.006	0.207	9.29	5.28	0.000	46	12	44
		0.001	0.025	81	0.015	0.006	0.207	9.24	5.26	0.000	3	-19	30
		0.005	0.120	35	0.090	0.008	0.268	8.92	5.18	0.000	3	-6	8

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 7.36, p = 0.000 (0.050)         Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.050)     FWHM = 17.8 17.1 15.3 mm mm mm; 11.9 11.4 10.2 {voxels}
Expected voxels per cluster, <k> = 12.466             Volume: 1676116 = 496627 voxels = 337.3 resels
Expected number of clusters, <c> = 0.05               Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1379.08 voxels)
FWEp: 7.364, FDRp: 11.634, FWEc: 1, FDRc: 81          Page 1
```

Figure E.1. Button press: brain glass image and statistic tables for uncorrected p-value < 0.001 at the left and FWE corrected p-value<0.05 at the right.

E.2) Level 3 over level 0 formulas related activations: brain glass image and statistic tables with uncorrected p-value < 0.001 for all formulas at the left and symbol formulas at the right.



```
SPMresults: ./group/contrast_13
Height threshold T = 3.732834  {p<0.001 (unc.)}
Extent threshold k = 0 voxels
```

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.329	10	0.102	0.128	507	0.013	0.300	0.444	6.06	4.24	0.000	-42	-64	19
		0.547	0.315	202	0.095	0.372	0.444	5.86	4.16	0.000	-50	-36	32
						0.954	0.852	4.48	3.51	0.000	-44	-46	36
		0.277	0.194	327	0.039	0.755	0.852	5.04	3.80	0.000	-52	-31	-16
						0.971	0.852	4.37	3.46	0.000	-58	-36	-24
		0.998	0.934	9	0.744	0.975	0.852	4.34	3.44	0.000	-4	44	58
		0.997	0.934	13	0.685	0.987	0.852	4.22	3.37	0.000	-36	-80	42
		0.982	0.934	35	0.481	0.988	0.852	4.20	3.36	0.000	-20	22	48
		0.998	0.934	8	0.760	0.993	0.853	4.13	3.32	0.000	-32	-32	-12
		0.999	0.934	3	0.867	0.995	0.853	4.06	3.28	0.001	-26	-46	-2
		0.996	0.934	14	0.672	0.999	0.899	3.91	3.20	0.001	-68	-38	-12
		1.000	0.934	1	0.934	1.000	0.912	3.80	3.13	0.001	-10	-14	67

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (1.000)           Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000)       FWHM = 15.9 15.4 14.5 mm mm mm; 10.6 10.2 9.7 {voxels}
Expected voxels per cluster, <k> = 74.170               Volume: 1676116 = 496627 voxels = 443.9 resels
Expected number of clusters, <c> = 8.36                 Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1047.94 voxels)
FWEp: 7.578, FDRp: Inf, FWEc: Inf, FDRc: Inf
```

```
SPMresults: ./group/contrast_19
Height threshold T = 3.732834  {p<0.001 (unc.)}
Extent threshold k = 0 voxels
```

Statistics: p-values adjusted for search volume

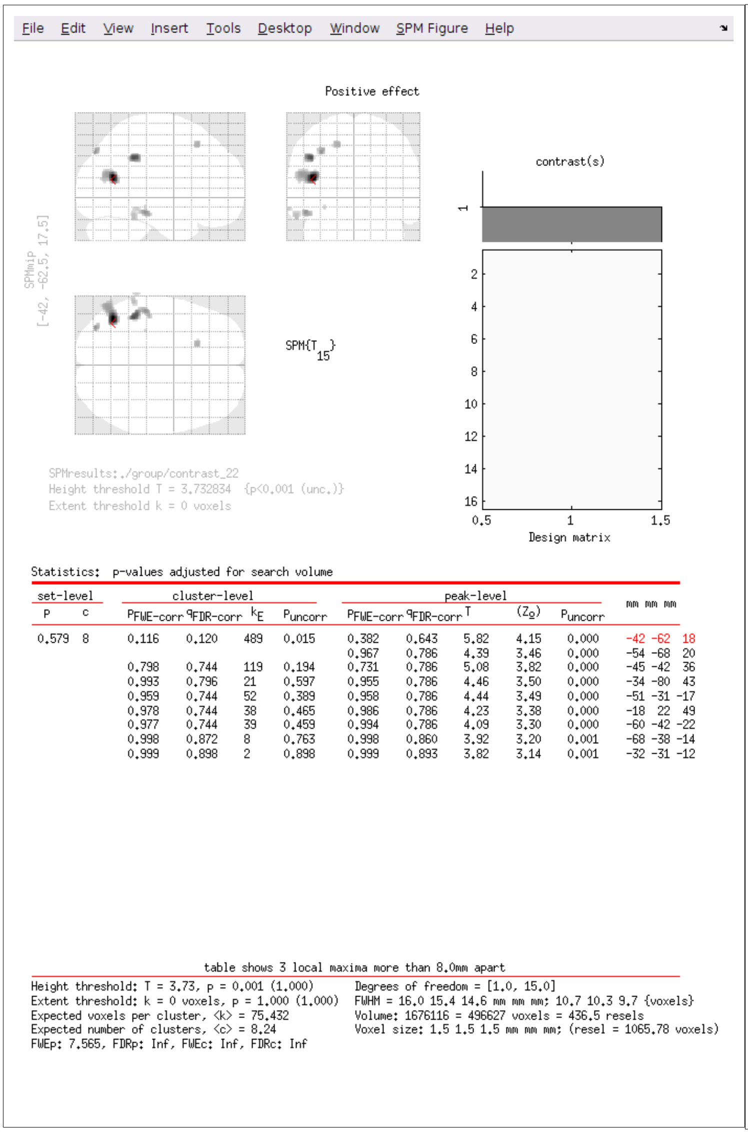
set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	18	0.110	0.305	572	0.017	0.375	0.485	5.70	4.09	0.000	-42	-61	19
						0.996	0.864	3.91	3.19	0.001	-52	-70	18
		0.941	0.797	58	0.412	0.441	0.485	5.54	4.03	0.000	-46	-30	-20
		0.892	0.797	83	0.324	0.512	0.485	5.39	3.96	0.000	-14	54	-12
						0.969	0.768	4.24	3.38	0.000	-24	52	-8
		0.793	0.797	124	0.229	0.512	0.485	5.38	3.96	0.000	-45	-1	-34
		0.675	0.797	169	0.164	0.664	0.485	5.08	3.82	0.000	52	-4	-32
						0.979	0.768	4.16	3.34	0.000	44	-2	-32
		0.805	0.797	119	0.239	0.671	0.485	5.07	3.81	0.000	-42	34	-10
		0.952	0.797	51	0.443	0.715	0.485	4.98	3.77	0.000	-62	-49	43
		0.231	0.345	409	0.038	0.738	0.485	4.93	3.74	0.000	-6	60	20
						0.982	0.768	4.14	3.33	0.000	-18	66	16
		0.904	0.797	77	0.343	0.872	0.603	4.62	3.59	0.000	12	-14	74
		0.950	0.797	52	0.438	0.883	0.603	4.59	3.57	0.000	32	-32	20
		0.979	0.847	30	0.565	0.949	0.758	4.35	3.45	0.000	-4	42	-11
		0.978	0.847	31	0.558	0.956	0.758	4.32	3.43	0.000	-6	50	-0
		0.993	0.936	12	0.733	0.981	0.768	4.14	3.33	0.000	-22	8	-5
		0.998	0.936	3	0.884	0.990	0.831	4.04	3.27	0.001	-20	-20	38
		0.998	0.936	3	0.884	0.995	0.864	3.93	3.21	0.001	40	-30	68
		0.993	0.936	12	0.733	0.995	0.864	3.92	3.20	0.001	-62	-28	-4
		0.997	0.936	4	0.861	0.998	0.953	3.80	3.13	0.001	-69	-24	-6
		0.998	0.943	1	0.943	0.999	0.964	3.76	3.11	0.001	-44	-36	28

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.999)           Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.999)       FWHM = 17.1 16.5 15.6 mm mm mm; 11.4 11.0 10.4 {voxels}
Expected voxels per cluster, <k> = 92.293               Volume: 1676116 = 496627 voxels = 356.7 resels
Expected number of clusters, <c> = 6.86                 Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1304.00 voxels)
FWEp: 7.407, FDRp: Inf, FWEc: Inf, FDRc: Inf
```

Figure E.2. Level 3 over level 0 formulas related activations: brain glass image and statistic tables with uncorrected p-value < 0.001 for all formulas at the left and symbols at the right.

E.3) Linear complexity contrast: brain glass image and statistic tables with uncorrected p-value < 0.001 for all formulas at the left and digit formulas at the right.



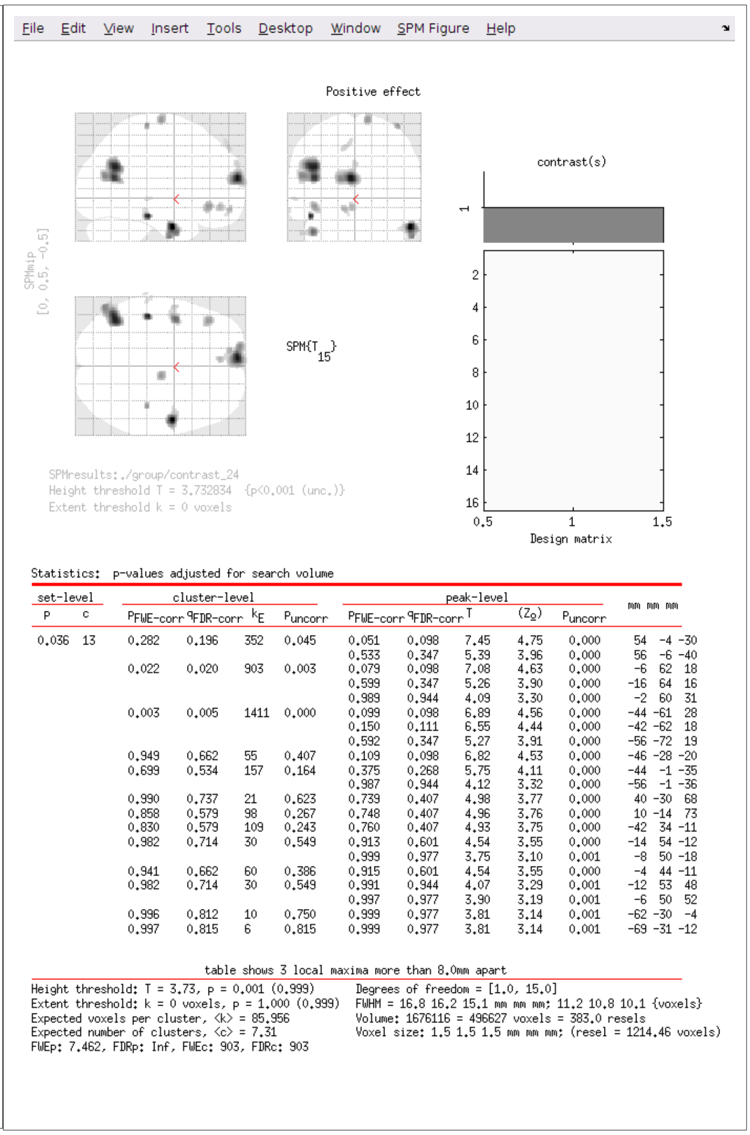
SPMresults:./group/contrast_22
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}	mm	mm	mm
0.579	8	0.116	0.120	489	0.015	0.382	0.643	5.82	4.15	0.000	-42	-62	18
						0.967	0.786	4.39	3.46	0.000	-54	-68	20
		0.798	0.744	119	0.194	0.731	0.786	5.08	3.82	0.000	-45	-42	36
		0.993	0.796	21	0.597	0.955	0.786	4.46	3.50	0.000	-34	-80	43
		0.959	0.744	52	0.389	0.958	0.786	4.44	3.49	0.000	-51	-31	-17
		0.978	0.744	38	0.465	0.986	0.786	4.23	3.38	0.000	-18	22	49
		0.977	0.744	39	0.459	0.994	0.786	4.09	3.30	0.000	-60	-42	-22
		0.998	0.872	8	0.763	0.998	0.860	3.92	3.20	0.001	-68	-38	-14
		0.999	0.898	2	0.898	0.999	0.893	3.82	3.14	0.001	-32	-31	-12

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (1.000)          Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000)      FWHM = 16.0 15.4 14.6 mm mm mm; 10.7 10.3 9.7 {voxels}
Expected voxels per cluster, <k> = 75.432              Volume: 1676116 = 496627 voxels = 436.5 resels
Expected number of clusters, <c> = 8.24                Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1065.78 voxels)
FWEp: 7.565, FDRp: Inf, FWEc: Inf, FDRc: Inf
```



SPMresults:./group/contrast_24
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

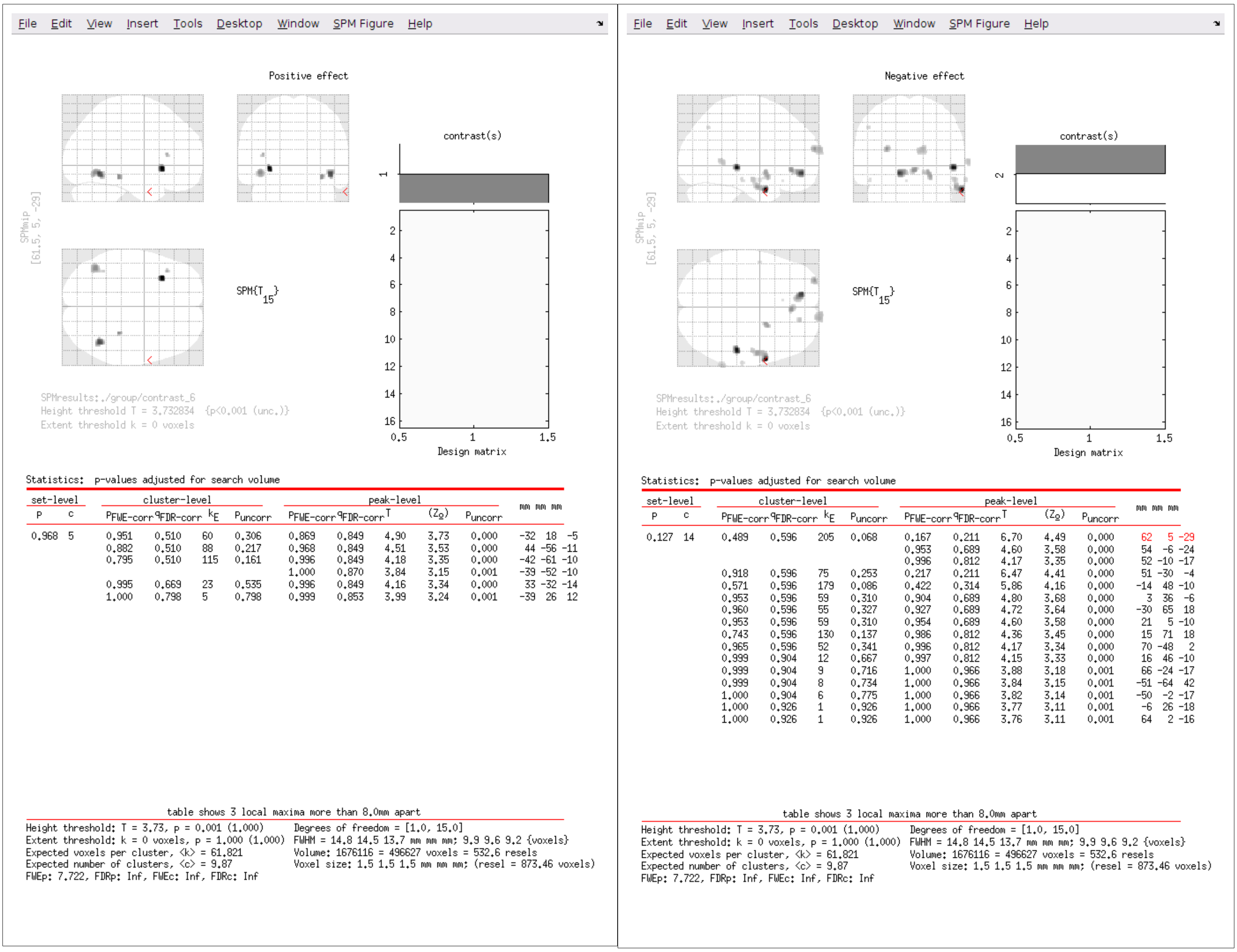
set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}	mm	mm	mm
0.036	13	0.282	0.196	352	0.045	0.051	0.098	7.45	4.75	0.000	54	-4	-30
						0.533	0.347	5.39	3.96	0.000	56	-6	-40
		0.022	0.020	903	0.003	0.079	0.098	7.08	4.63	0.000	-6	62	18
						0.599	0.347	5.26	3.90	0.000	-16	64	16
						0.989	0.944	4.09	3.30	0.000	-2	60	31
		0.003	0.005	1411	0.000	0.099	0.098	6.89	4.56	0.000	-44	-61	28
						0.150	0.111	6.55	4.44	0.000	-42	-62	18
						0.592	0.347	5.27	3.91	0.000	-56	-72	19
		0.949	0.662	55	0.407	0.109	0.098	6.82	4.53	0.000	-46	-28	-20
		0.699	0.534	157	0.164	0.375	0.268	5.75	4.11	0.000	-44	-1	-35
						0.987	0.944	4.12	3.32	0.000	-56	-1	-36
		0.990	0.737	21	0.623	0.739	0.407	4.98	3.77	0.000	40	-30	68
		0.858	0.579	98	0.267	0.748	0.407	4.96	3.76	0.000	10	-14	73
		0.830	0.579	109	0.243	0.760	0.407	4.93	3.75	0.000	-42	34	-11
		0.982	0.714	30	0.549	0.913	0.601	4.54	3.55	0.000	-14	54	-12
						0.999	0.977	3.75	3.10	0.001	-8	50	-18
		0.941	0.662	60	0.386	0.915	0.601	4.54	3.55	0.000	-4	44	-11
		0.982	0.714	30	0.549	0.991	0.944	4.07	3.29	0.001	-12	53	48
						0.997	0.977	3.90	3.19	0.001	-6	50	52
		0.996	0.812	10	0.750	0.999	0.977	3.81	3.14	0.001	-62	-30	-4
		0.997	0.815	6	0.815	0.999	0.977	3.81	3.14	0.001	-69	-31	-12

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.999)          Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.999)      FWHM = 16.8 16.2 15.1 mm mm mm; 11.2 10.8 10.1 {voxels}
Expected voxels per cluster, <k> = 85.956              Volume: 1676116 = 496627 voxels = 383.0 resels
Expected number of clusters, <c> = 7.31                Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1214.46 voxels)
FWEp: 7.462, FDRp: Inf, FWEc: 903, FDRc: 903
```

Figure E.3. Linear complexity contrast: brain glass image and statistic tables with uncorrected p-value < 0.001 for all formulas at the left and digit formulas at the right.

E.4) Symbol vs digit formulas contrast: brain glass image and statistic tables with uncorrected p-value < 0.001 for symbol over digit formulas at the left and digit over symbol formulas at the right.



SPMresults:./group/contrast_6
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}	mm	mm	mm
0.968	5	0.951	0.510	60	0.306	0.869	0.849	4.90	3.73	0.000	-32	18	-5
		0.882	0.510	88	0.217	0.968	0.849	4.51	3.53	0.000	44	-56	-11
		0.795	0.510	115	0.161	0.996	0.849	4.18	3.35	0.000	-42	-61	-10
						1.000	0.870	3.84	3.15	0.001	-39	-52	-10
		0.995	0.669	23	0.535	0.996	0.849	4.16	3.34	0.000	33	-32	-14
		1.000	0.798	5	0.798	0.999	0.853	3.99	3.24	0.001	-39	26	12

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 1.000 (1.000)
Expected voxels per cluster, <k> = 61.821
Expected number of clusters, <c> = 9.87
FWEp: 7.722, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 14.8 14.5 13.7 mm mm mm; 9.9 9.6 9.2 {voxels}
Volume: 1676116 = 496627 voxels = 532.6 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 873.46 voxels)

SPMresults:./group/contrast_6
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}	mm	mm	mm
0.127	14	0.489	0.596	205	0.068	0.167	0.211	6.70	4.49	0.000	62	5	-29
						0.953	0.689	4.60	3.58	0.000	54	-6	-24
						0.996	0.812	4.17	3.35	0.000	52	-10	-17
		0.918	0.596	75	0.253	0.217	0.211	6.47	4.41	0.000	51	-30	-4
		0.571	0.596	179	0.086	0.422	0.314	5.86	4.16	0.000	-14	48	-10
		0.953	0.596	59	0.310	0.904	0.689	4.80	3.68	0.000	3	36	-6
		0.960	0.596	55	0.327	0.927	0.689	4.72	3.64	0.000	-30	65	18
		0.953	0.596	59	0.310	0.954	0.689	4.60	3.58	0.000	21	5	-10
		0.743	0.596	130	0.137	0.986	0.812	4.36	3.45	0.000	15	71	18
		0.965	0.596	52	0.341	0.996	0.812	4.17	3.34	0.000	70	-48	2
		0.999	0.904	12	0.667	0.997	0.812	4.15	3.33	0.000	16	46	-10
		0.999	0.904	9	0.716	1.000	0.966	3.88	3.18	0.001	66	-24	-17
		0.999	0.904	8	0.734	1.000	0.966	3.84	3.15	0.001	-51	-64	42
		1.000	0.904	6	0.775	1.000	0.966	3.82	3.14	0.001	-50	-2	-17
		1.000	0.926	1	0.926	1.000	0.966	3.77	3.11	0.001	-6	26	-18
		1.000	0.926	1	0.926	1.000	0.966	3.76	3.11	0.001	64	2	-16

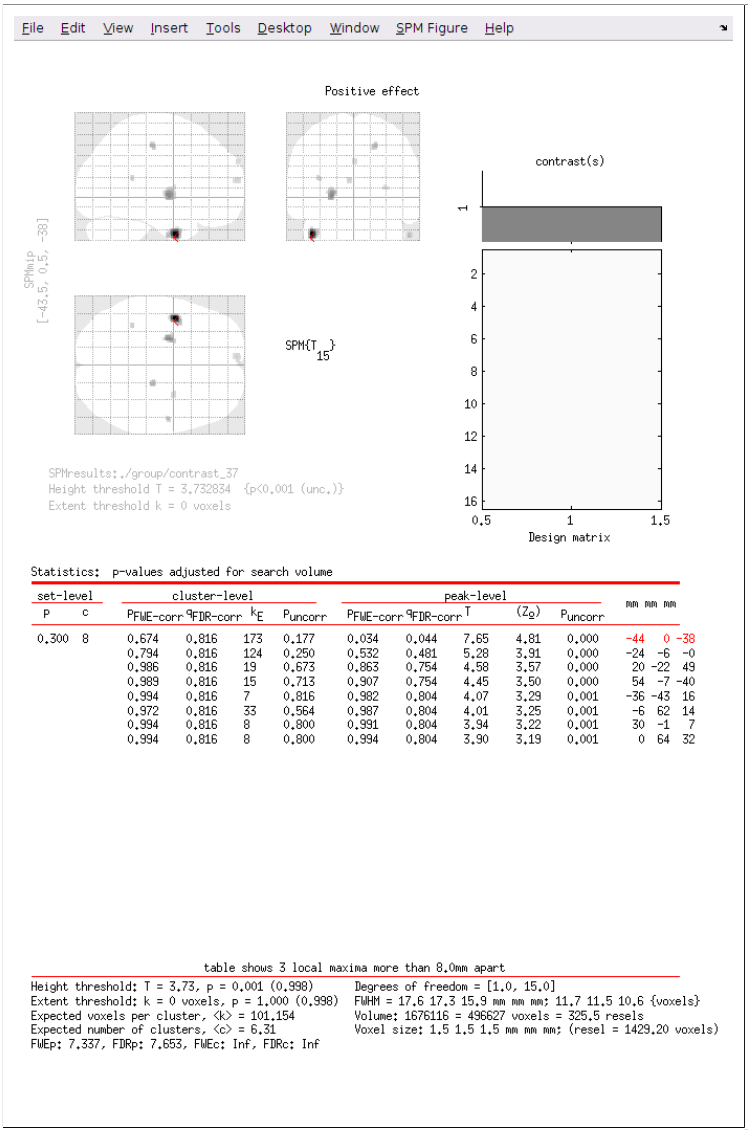
table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 1.000 (1.000)
Expected voxels per cluster, <k> = 61.821
Expected number of clusters, <c> = 9.87
FWEp: 7.722, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 14.8 14.5 13.7 mm mm mm; 9.9 9.6 9.2 {voxels}
Volume: 1676116 = 496627 voxels = 532.6 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 873.46 voxels)

Figure E.4. Symbol vs digit formulas contrast: brain glass image and statistic tables with uncorrected p-value < 0.001 for symbol over digit formulas at the left and digit over symbol formulas at the right.

E.5) linear complexity in symbol formulas over linear complexity in digit formulas: brain glass image and statistic tables with uncorrected p-value < 0.001 at the left and for FWE corrected p-value < 0.05 at the right.



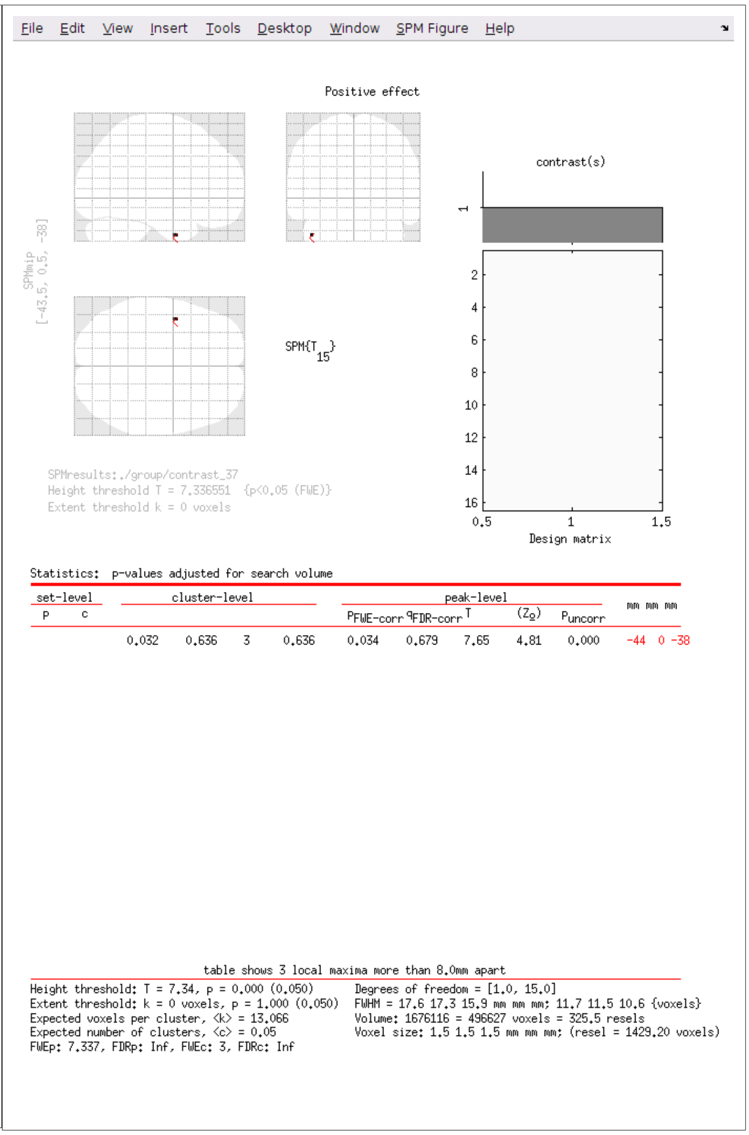
SPMresults: ./group/contrast_37
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.300	8	0.674	0.816	173	0.177	0.034	0.044	7.65	4.81	0.000	-44	0	-38
		0.794	0.816	124	0.250	0.532	0.481	5.28	3.91	0.000	-24	-6	-0
		0.986	0.816	19	0.673	0.863	0.754	4.58	3.57	0.000	20	-22	49
		0.989	0.816	15	0.713	0.907	0.754	4.45	3.50	0.000	54	-7	-40
		0.994	0.816	7	0.816	0.982	0.804	4.07	3.29	0.001	-36	-43	16
		0.972	0.816	33	0.564	0.987	0.804	4.01	3.25	0.001	-6	62	14
		0.994	0.816	8	0.800	0.991	0.804	3.94	3.22	0.001	30	-1	7
		0.994	0.816	8	0.800	0.994	0.804	3.90	3.19	0.001	0	64	32

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.998)          Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.998)      FWHM = 17.6 17.3 15.9 mm mm mm; 11.7 11.5 10.6 {voxels}
Expected voxels per cluster, <k> = 101.154             Volume: 1676116 = 496627 voxels = 325.5 resels
Expected number of clusters, <c> = 6.31                Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1429.20 voxels)
FWEp: 7.337, FDRp: 7.653, FWEc: Inf, FDRc: Inf
```



SPMresults: ./group/contrast_37
Height threshold T = 7.336551 {p<0.05 (FWE)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c					$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
		0.032	0.636	3	0.636	0.034	0.679	7.65	4.81	0.000	-44	0	-38

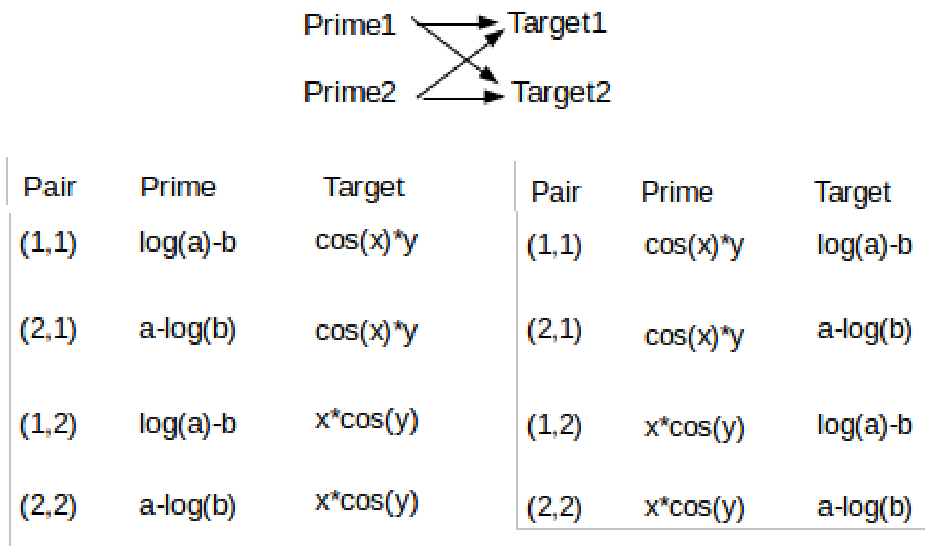
table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 7.34, p = 0.000 (0.050)          Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.050)      FWHM = 17.6 17.3 15.9 mm mm mm; 11.7 11.5 10.6 {voxels}
Expected voxels per cluster, <k> = 13.066              Volume: 1676116 = 496627 voxels = 325.5 resels
Expected number of clusters, <c> = 0.05                Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1429.20 voxels)
FWEp: 7.337, FDRp: Inf, FWEc: 3, FDRc: Inf
```

Figure E.5. linear complexity in symbol formulas over linear complexity in digit formulas: brain glass image and statistic tables with uncorrected p-value < 0.001 at the left and for FWE corrected p-value < 0.05 at the right.

Appendix F. Priming Trees Stimuli

F.1) Diagram of study of priming effect



Pair	Prime	Target
(1,1)	log(a)-b	cos(x)*y
(2,1)	a-log(b)	cos(x)*y
(1,2)	log(a)-b	x*cos(y)
(2,2)	a-log(b)	x*cos(y)

Pair	Prime	Target
(1,1)	cos(x)*y	log(a)-b
(2,1)	cos(x)*y	a-log(b)
(1,2)	x*cos(y)	log(a)-b
(2,2)	x*cos(y)	a-log(b)

Figure F.1. Diagram of complete set of stimuli for a pair of formula structures to effectively study a priming effect. At the top a small diagram showing the necessary interactions to study. At the left the combinations of stimuli as prime and target to study the effect. At the right exactly the same combinations but with a swap in the prime and target roles to have the same population of stimuli in both roles for robustness in estimation.

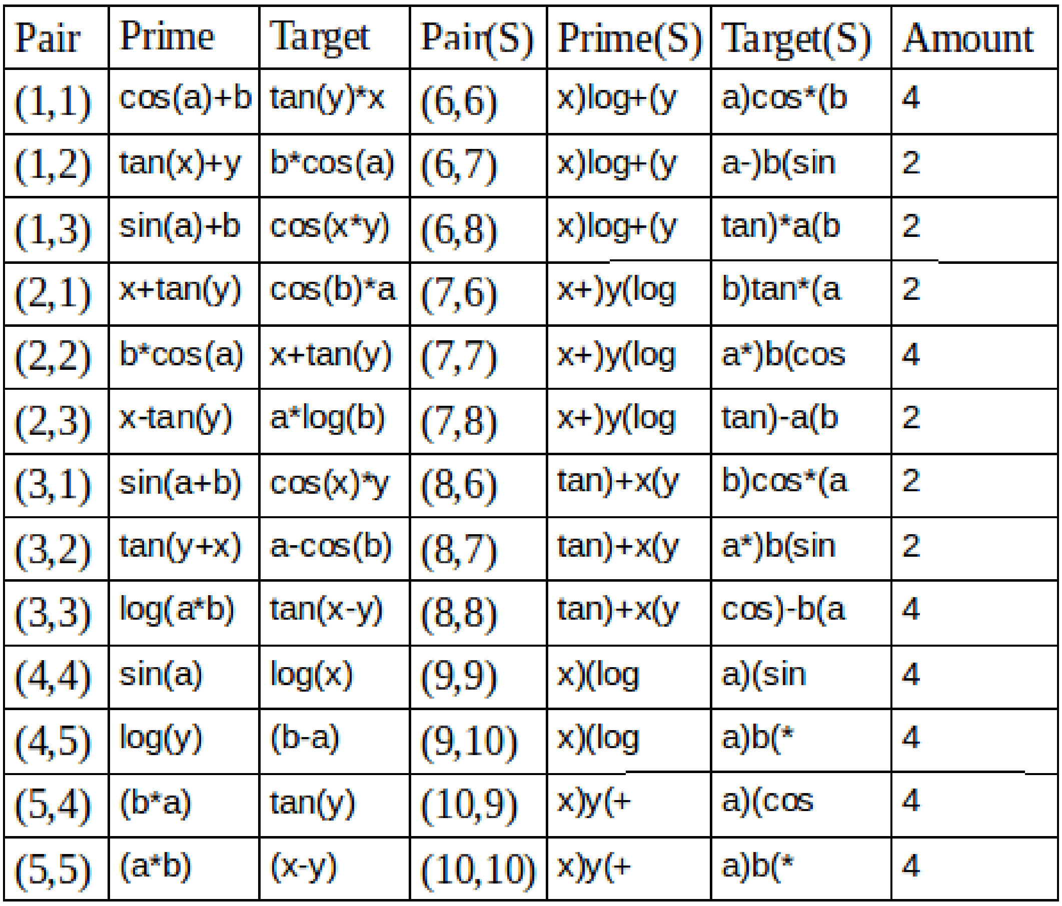
F.2) Formula patterns

Formula type	Example	Structure	#(Op)	#(Var)	#(char)
Una left (1)	log(y)+x	una(var)binvar	2	2	8
Una right (2)	x*sin(y)	varbinuna(var)	2	2	8
Una top (3)	tan(a-b)	una(varbinvar)	2	2	8
Only una (4)	tan(y)	una(var)	1	1	6
Only bin (5)	(y+x)	(varbinvar)	1	2	5
Scrambled large 1 (6)	x)log+(y	var)ulbin(var	2	2	8
Scrambled large 2 (7)	x+)y(log	varbin)var(ul	2	2	8
Scrambled large 3 (8)	tan)+x(y	ul)binvar(var	2	2	8
Scrambled una (9)	x)(log	var)(ul	1	1	6
Scrambled bin (10)	x)y(+	var)var(bin	1	2	5

Table F.1. Table of stimuli examples and characteristics for all formula patterns

F.3) Trial design extra details

80 valid formula pairs comprise the stimuli for a given run of acquisition, alongside 10 formula pairs with an “@” to perform the task. Out of the 80 formula pairs, 40 correspond to structured patterns and 40 to scrambled patterns. More details on examples and the proportion of appearance of each pair for the structured and scrambled case can be seen in the following table.



Pair	Prime	Target	Pair(S)	Prime(S)	Target(S)	Amount
(1,1)	cos(a)+b	tan(y)*x	(6,6)	x)log+(y	a)cos*(b	4
(1,2)	tan(x)+y	b*cos(a)	(6,7)	x)log+(y	a-)b(sin	2
(1,3)	sin(a)+b	cos(x*y)	(6,8)	x)log+(y	tan)*a(b	2
(2,1)	x+tan(y)	cos(b)*a	(7,6)	x+)y(log	b)tan*(a	2
(2,2)	b*cos(a)	x+tan(y)	(7,7)	x+)y(log	a*)b(cos	4
(2,3)	x-tan(y)	a*log(b)	(7,8)	x+)y(log	tan)-a(b	2
(3,1)	sin(a+b)	cos(x)*y	(8,6)	tan)+x(y	b)cos*(a	2
(3,2)	tan(y+x)	a-cos(b)	(8,7)	tan)+x(y	a*)b(sin	2
(3,3)	log(a*b)	tan(x-y)	(8,8)	tan)+x(y	cos)-b(a	4
(4,4)	sin(a)	log(x)	(9,9)	x)(log	a)(sin	4
(4,5)	log(y)	(b-a)	(9,10)	x)(log	a)b(*	4
(5,4)	(b*a)	tan(y)	(10,9)	x)y(+	a)(cos	4
(5,5)	(a*b)	(x-y)	(10,10)	x)y(+	a)b(*	4

Table F.2. Table of stimuli examples for all formula pairs presented, with their respective proportions of presentation in a run of acquisition for a total of 80 stimuli. Scrambled paris are denoted by (S). The amount applies equally for both structured and scrambled pairs. All the table gives a total of 40 stimuli for each type of pattern. Giving 80 stimuli in total. An extra 10 stimuli modified for the task would add up to 90 stimuli per run of acquisition.

Appendix G. Priming Trees Behavioral results

G.1) Task accuracy

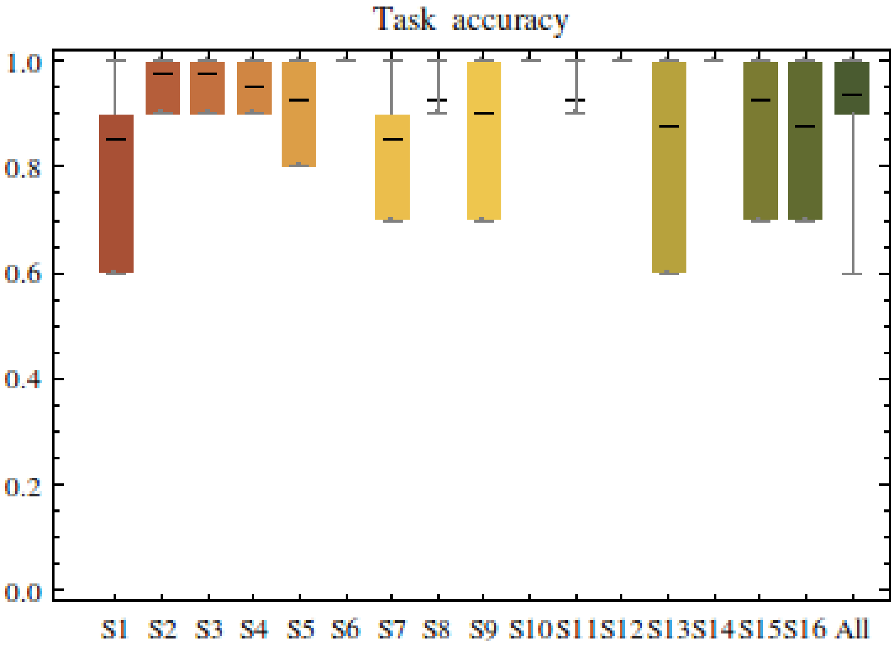


Figure G.1. Accuracy of task in Priming Trees

G.2) Task false positive rate

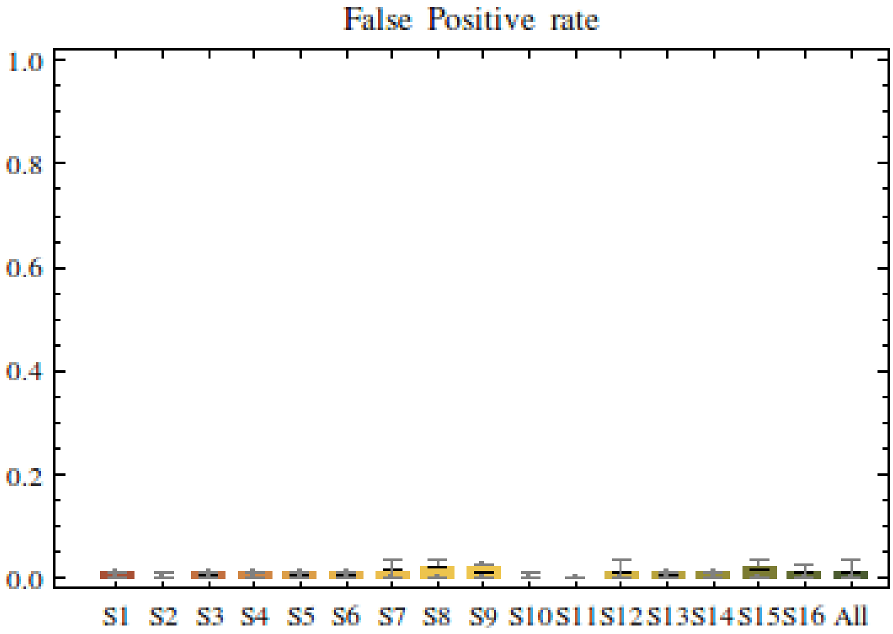
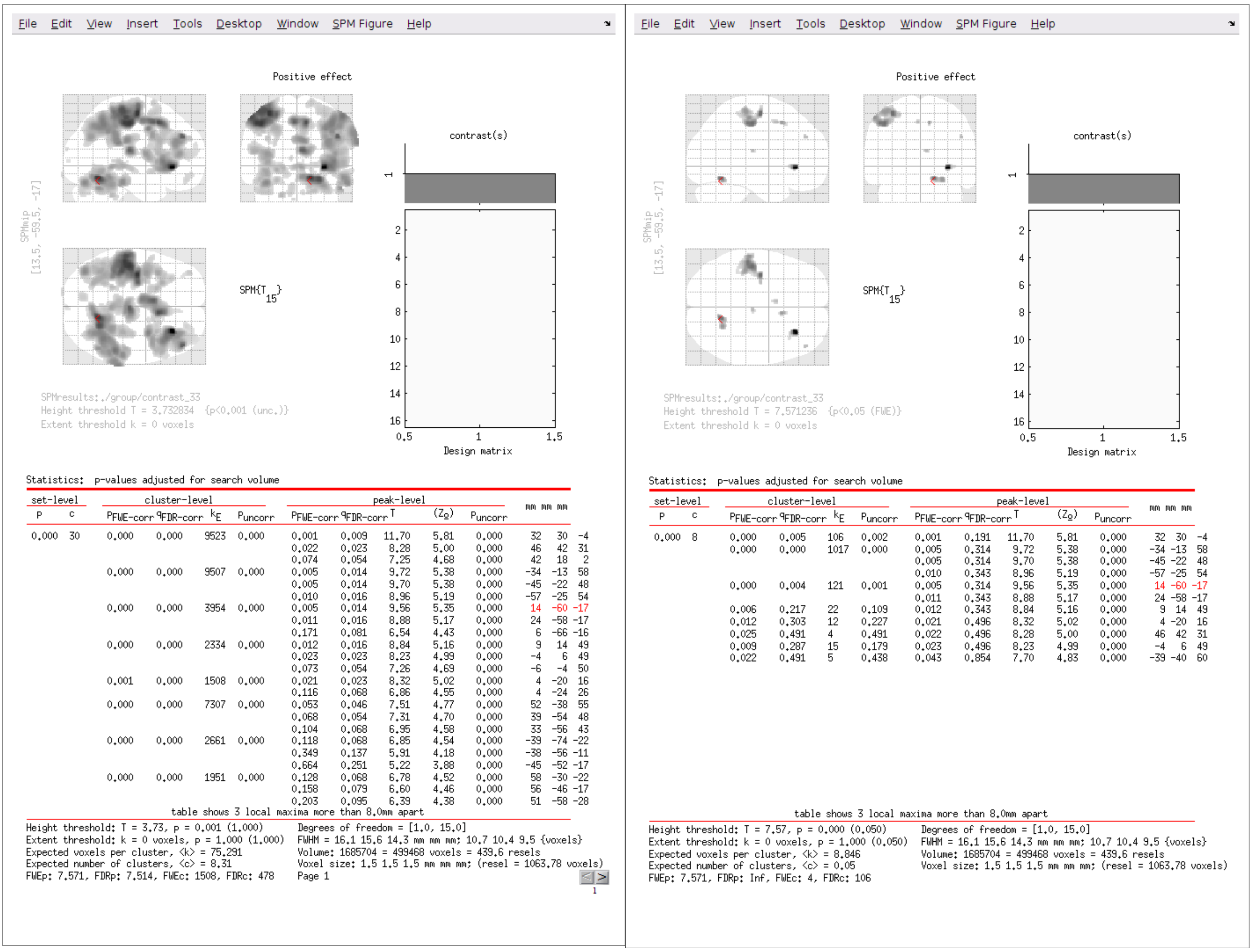


Figure G.2. False positive rate of task in Priming Trees

Appendix H. Priming Trees Brain Activations

H.1) Button press: brain glass image and statistic tables for uncorrected p-value < 0.001 at the left and FWE corrected p-value<0.05 at the right.



SPMresults:./group/contrast_33
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level					mm	mm	mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	($Z_{\equiv}$)	p_{uncorr}			
0.000	30	0.000	0.000	9523	0.000	0.001	0.009	11.70	5.81	0.000	32	30	-4
						0.022	0.023	8.28	5.00	0.000	46	42	31
						0.074	0.054	7.25	4.68	0.000	42	18	2
		0.000	0.000	9507	0.000	0.005	0.014	9.72	5.38	0.000	-34	-13	58
						0.005	0.014	9.70	5.38	0.000	-45	-22	48
						0.010	0.016	8.96	5.19	0.000	-57	-25	54
		0.000	0.000	3954	0.000	0.005	0.014	9.56	5.35	0.000	14	-60	-17
						0.011	0.016	8.88	5.17	0.000	24	-58	-17
						0.171	0.081	6.54	4.43	0.000	6	-66	-16
		0.000	0.000	2334	0.000	0.012	0.016	8.84	5.16	0.000	9	14	49
						0.023	0.023	8.23	4.99	0.000	-4	6	49
						0.073	0.054	7.26	4.69	0.000	-6	-4	50
		0.001	0.000	1508	0.000	0.021	0.023	8.32	5.02	0.000	4	-20	16
						0.116	0.068	6.86	4.55	0.000	4	-24	26
		0.000	0.000	7307	0.000	0.053	0.046	7.51	4.77	0.000	52	-38	55
						0.068	0.054	7.31	4.70	0.000	39	-54	48
						0.104	0.068	6.95	4.58	0.000	33	-56	43
		0.000	0.000	2661	0.000	0.118	0.068	6.85	4.54	0.000	-39	-74	-22
						0.349	0.137	5.91	4.18	0.000	-38	-56	-11
						0.664	0.251	5.22	3.88	0.000	-45	-52	-17
		0.000	0.000	1951	0.000	0.128	0.068	6.78	4.52	0.000	58	-30	-22
						0.158	0.079	6.60	4.46	0.000	56	-46	-17
						0.203	0.095	6.39	4.38	0.000	51	-58	-28

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (1.000)        Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000)    FWHM = 16.1 15.6 14.3 mm mm mm; 10.7 10.4 9.5 {voxels}
Expected voxels per cluster, <k> = 75.291            Volume: 1685704 = 499468 voxels = 439.6 resels
Expected number of clusters, <c> = 8.31              Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1063.78 voxels)
FWEp: 7.571, FDRp: 7.514, FWEc: 1508, FDRc: 478      Page 1
```

SPMresults:./group/contrast_33
Height threshold T = 7.571236 {p<0.05 (FWE)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

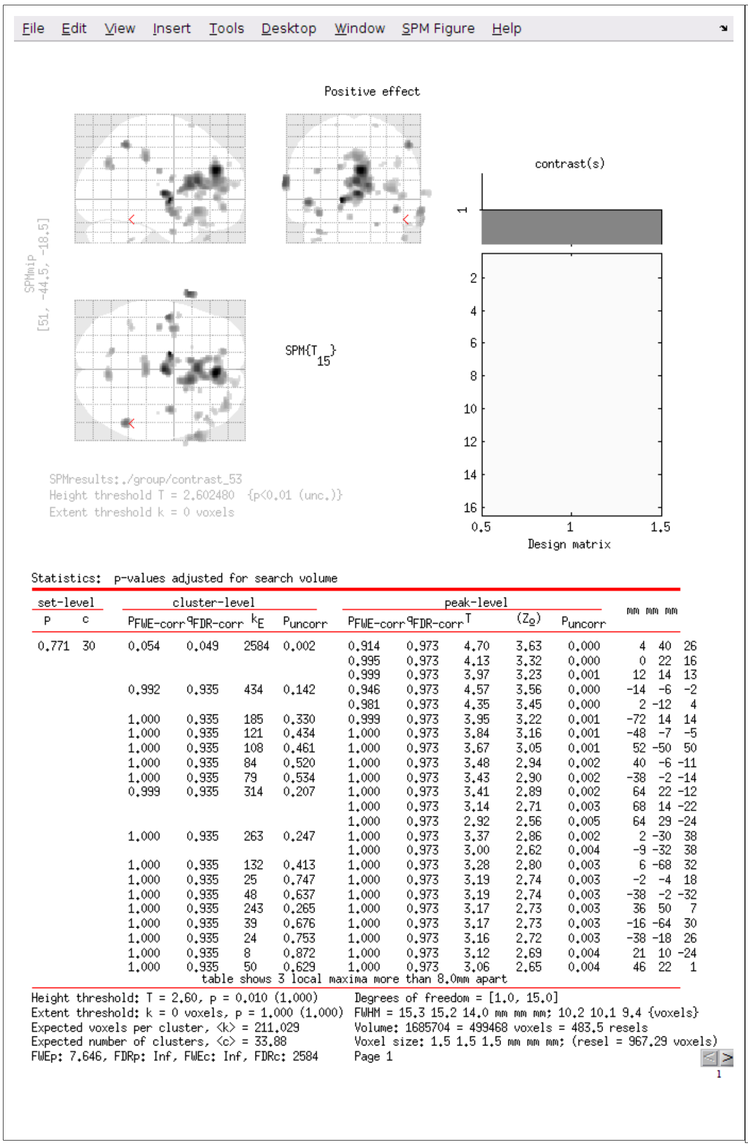
set-level		cluster-level				peak-level					mm	mm	mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	($Z_{\equiv}$)	p_{uncorr}			
0.000	8	0.000	0.005	106	0.002	0.001	0.191	11.70	5.81	0.000	32	30	-4
		0.000	0.000	1017	0.000	0.005	0.314	9.72	5.38	0.000	-34	-13	58
						0.005	0.314	9.70	5.38	0.000	-45	-22	48
						0.010	0.343	8.96	5.19	0.000	-57	-25	54
		0.000	0.004	121	0.001	0.005	0.314	9.56	5.35	0.000	14	-60	-17
						0.011	0.343	8.88	5.17	0.000	24	-58	-17
		0.006	0.217	22	0.109	0.012	0.343	8.84	5.16	0.000	9	14	49
		0.012	0.303	12	0.227	0.021	0.496	8.32	5.02	0.000	4	-20	16
		0.025	0.491	4	0.491	0.022	0.496	8.28	5.00	0.000	46	42	31
		0.009	0.287	15	0.179	0.023	0.496	8.23	4.99	0.000	-4	6	49
		0.022	0.491	5	0.438	0.043	0.854	7.70	4.83	0.000	-39	-40	60

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 7.57, p = 0.000 (0.050)        Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.050)    FWHM = 16.1 15.6 14.3 mm mm mm; 10.7 10.4 9.5 {voxels}
Expected voxels per cluster, <k> = 8.846             Volume: 1685704 = 499468 voxels = 439.6 resels
Expected number of clusters, <c> = 0.05              Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1063.78 voxels)
FWEp: 7.571, FDRp: Inf, FWEc: 4, FDRc: 106
```

Figure H.1. Button press: brain glass image and statistic tables for uncorrected p-value < 0.001 at the left and FWE corrected p-value<0.05 at the right.

H.2) Complexity: brain glass image and statistic tables for uncorrected p-value < 0.001 of 2-operator formulas over 1-operator formula at the left and the opposite at the right.



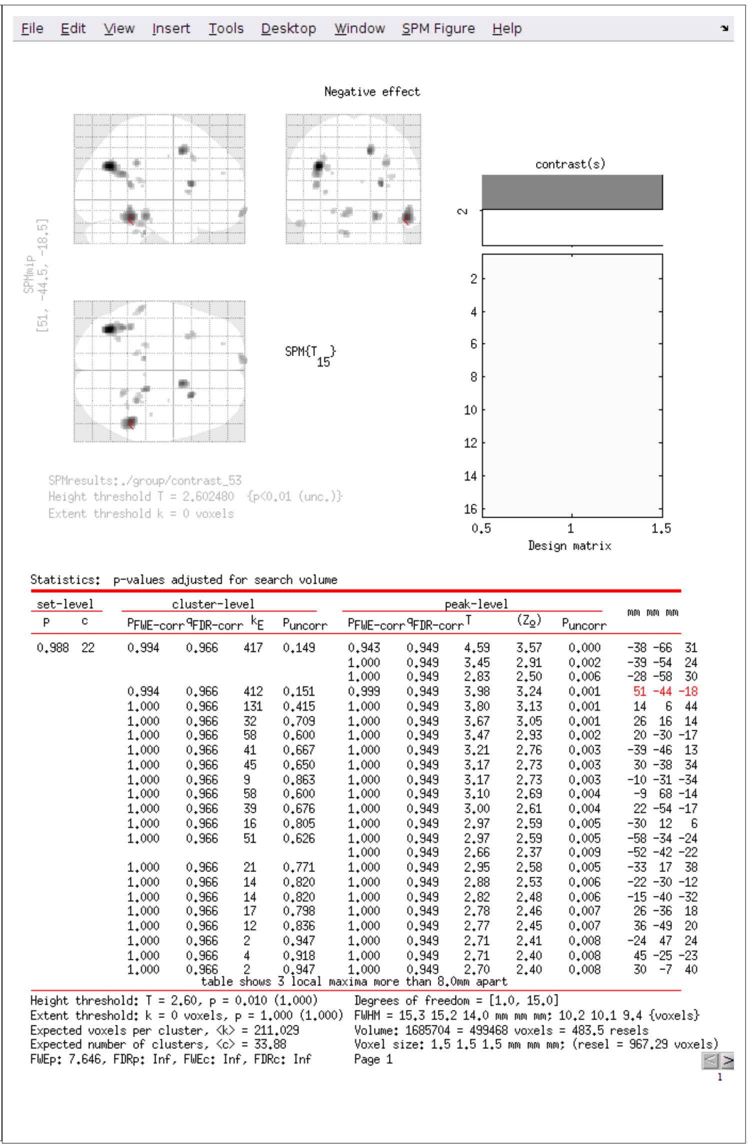
SPMresults:./group/contrast_53
Height threshold T = 2.602480 {p<0.01 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.771	30	0.054	0.049	2584	0.002	0.914	0.973	4.70	3.63	0.000	4	40	26
						0.995	0.973	4.13	3.32	0.000	0	22	16
						0.999	0.973	3.97	3.23	0.001	12	14	13
		0.992	0.935	434	0.142	0.946	0.973	4.57	3.56	0.000	-14	-6	-2
						0.981	0.973	4.35	3.45	0.000	2	-12	4
		1.000	0.935	185	0.330	0.999	0.973	3.95	3.22	0.001	-72	14	14
		1.000	0.935	121	0.434	1.000	0.973	3.84	3.16	0.001	-48	-7	-5
		1.000	0.935	108	0.461	1.000	0.973	3.67	3.05	0.001	52	-50	50
		1.000	0.935	84	0.520	1.000	0.973	3.48	2.94	0.002	40	-6	-11
		1.000	0.935	79	0.534	1.000	0.973	3.43	2.90	0.002	-38	-2	-14
		0.999	0.935	314	0.207	1.000	0.973	3.41	2.89	0.002	64	22	-12
						1.000	0.973	3.14	2.71	0.003	68	14	-22
						1.000	0.973	2.92	2.56	0.005	64	29	-24
		1.000	0.935	263	0.247	1.000	0.973	3.37	2.86	0.002	2	-30	38
						1.000	0.973	3.00	2.62	0.004	-9	-32	38
		1.000	0.935	132	0.413	1.000	0.973	3.28	2.80	0.003	6	-68	32
		1.000	0.935	25	0.747	1.000	0.973	3.19	2.74	0.003	-2	-4	18
		1.000	0.935	48	0.637	1.000	0.973	3.19	2.74	0.003	-38	-2	-32
		1.000	0.935	243	0.265	1.000	0.973	3.17	2.73	0.003	36	50	7
		1.000	0.935	39	0.676	1.000	0.973	3.17	2.73	0.003	-16	-64	30
		1.000	0.935	24	0.753	1.000	0.973	3.16	2.72	0.003	-38	-18	26
		1.000	0.935	8	0.872	1.000	0.973	3.12	2.69	0.004	21	10	-24
		1.000	0.935	50	0.629	1.000	0.973	3.06	2.65	0.004	46	22	1

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 2.60, p = 0.010 (1.000)       Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000)   FWHM = 15.3 15.2 14.0 mm mm mm; 10.2 10.1 9.4 {voxels}
Expected voxels per cluster, <k> = 211.029          Volume: 1685704 = 499468 voxels = 483.5 resels
Expected number of clusters, <c> = 33.88            Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 967.29 voxels)
FWEp: 7.646, FDRp: Inf, FWEc: Inf, FDRc: 2584       Page 1
```



SPMresults:./group/contrast_53
Height threshold T = 2.602480 {p<0.01 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

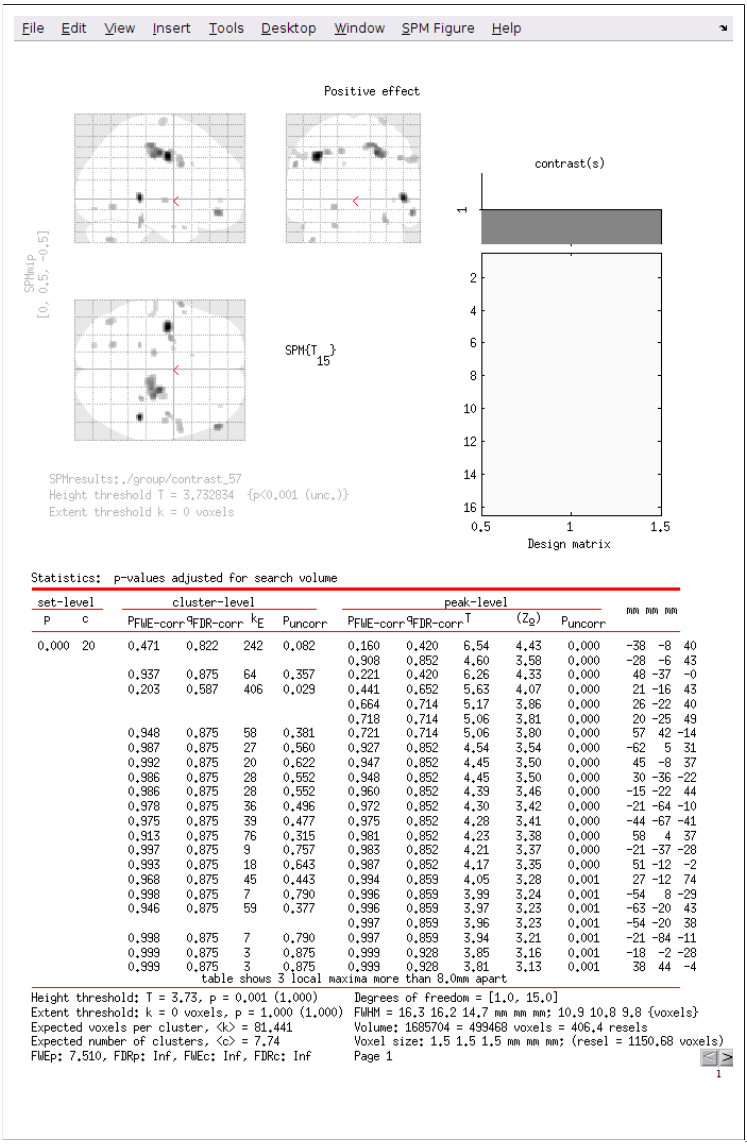
set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.988	22	0.994	0.966	417	0.149	0.943	0.949	4.59	3.57	0.000	-38	-66	31
						1.000	0.949	3.45	2.91	0.002	-39	-54	24
						1.000	0.949	2.83	2.50	0.006	-28	-58	30
		0.994	0.966	412	0.151	0.999	0.949	3.98	3.24	0.001	51	-44	-18
		1.000	0.966	131	0.415	1.000	0.949	3.80	3.13	0.001	14	6	44
		1.000	0.966	32	0.709	1.000	0.949	3.67	3.05	0.001	26	16	14
		1.000	0.966	58	0.600	1.000	0.949	3.47	2.93	0.002	20	-30	-17
		1.000	0.966	41	0.667	1.000	0.949	3.21	2.76	0.003	-39	-46	13
		1.000	0.966	45	0.650	1.000	0.949	3.17	2.73	0.003	30	-38	34
		1.000	0.966	9	0.863	1.000	0.949	3.17	2.73	0.003	-10	-31	-34
		1.000	0.966	58	0.600	1.000	0.949	3.10	2.69	0.004	-9	68	-14
		1.000	0.966	39	0.676	1.000	0.949	3.00	2.61	0.004	22	-54	-17
		1.000	0.966	16	0.805	1.000	0.949	2.97	2.59	0.005	-30	12	6
		1.000	0.966	51	0.626	1.000	0.949	2.97	2.59	0.005	-58	-34	-24
						1.000	0.949	2.66	2.37	0.009	-52	-42	-22
		1.000	0.966	21	0.771	1.000	0.949	2.95	2.58	0.005	-33	17	38
		1.000	0.966	14	0.820	1.000	0.949	2.88	2.53	0.006	-22	-30	-12
		1.000	0.966	14	0.820	1.000	0.949	2.82	2.48	0.006	-15	-40	-32
		1.000	0.966	17	0.798	1.000	0.949	2.78	2.46	0.007	26	-36	18
		1.000	0.966	12	0.836	1.000	0.949	2.77	2.45	0.007	36	-49	20
		1.000	0.966	2	0.947	1.000	0.949	2.71	2.41	0.008	-24	47	24
		1.000	0.966	4	0.918	1.000	0.949	2.71	2.40	0.008	45	-25	-23
		1.000	0.966	2	0.947	1.000	0.949	2.70	2.40	0.008	30	-7	40

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 2.60, p = 0.010 (1.000)       Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000)   FWHM = 15.3 15.2 14.0 mm mm mm; 10.2 10.1 9.4 {voxels}
Expected voxels per cluster, <k> = 211.029          Volume: 1685704 = 499468 voxels = 483.5 resels
Expected number of clusters, <c> = 33.88            Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 967.29 voxels)
FWEp: 7.646, FDRp: Inf, FWEc: Inf, FDRc: Inf        Page 1
```

Figure H.1. Complexity: brain glass image and statistic tables for uncorrected p-value < 0.001 of 2-operator formulas over 1-operator formula at the left and the opposite at the right.

H.3) Syntactic effects: brain glass image and statistic tables for uncorrected p-value < 0.001 of pairs of syntactically different formulas over pairs of syntactically similar formulas at the left and the same contrast including interaction with structure over non structure of formulas on the right.



SPMresults: ./group/contrast_57
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

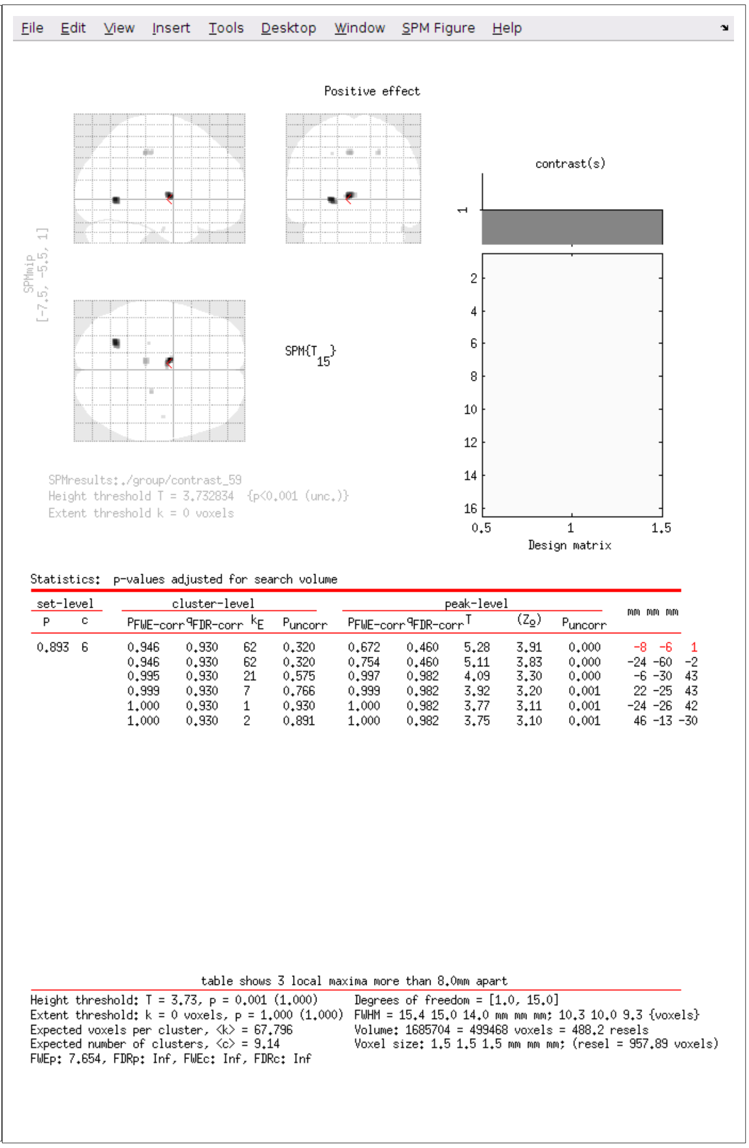
Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	20	0.471	0.822	242	0.082	0.160	0.420	6.54	4.43	0.000	-38	-8	40
						0.908	0.852	4.60	3.58	0.000	-28	-6	43
		0.937	0.875	64	0.357	0.221	0.420	6.26	4.33	0.000	48	-37	-0
		0.203	0.587	406	0.029	0.441	0.652	5.63	4.07	0.000	21	-16	43
						0.664	0.714	5.17	3.86	0.000	26	-22	40
						0.718	0.714	5.06	3.81	0.000	20	-25	49
		0.948	0.875	58	0.381	0.721	0.714	5.06	3.80	0.000	57	42	-14
		0.987	0.875	27	0.560	0.927	0.852	4.54	3.54	0.000	-62	5	31
		0.992	0.875	20	0.622	0.947	0.852	4.45	3.50	0.000	45	-8	37
		0.986	0.875	28	0.552	0.948	0.852	4.45	3.50	0.000	30	-36	-22
		0.986	0.875	28	0.552	0.960	0.852	4.39	3.46	0.000	-15	-22	44
		0.978	0.875	36	0.496	0.972	0.852	4.30	3.42	0.000	-21	-64	-10
		0.975	0.875	39	0.477	0.975	0.852	4.28	3.41	0.000	-44	-67	-41
		0.913	0.875	76	0.315	0.981	0.852	4.23	3.38	0.000	58	4	37
		0.997	0.875	9	0.757	0.983	0.852	4.21	3.37	0.000	-21	-37	-28
		0.993	0.875	18	0.643	0.987	0.852	4.17	3.35	0.000	51	-12	-2
		0.968	0.875	45	0.443	0.994	0.859	4.05	3.28	0.001	27	-12	74
		0.998	0.875	7	0.790	0.996	0.859	3.99	3.24	0.001	-54	8	-29
		0.946	0.875	59	0.377	0.996	0.859	3.97	3.23	0.001	-63	-20	43
						0.997	0.859	3.96	3.23	0.001	-54	-20	38
		0.998	0.875	7	0.790	0.997	0.859	3.94	3.21	0.001	-21	-84	-11
		0.999	0.875	3	0.875	0.999	0.928	3.85	3.16	0.001	-18	-2	-28
		0.999	0.875	3	0.875	0.999	0.928	3.81	3.13	0.001	38	44	-4

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 1.000 (1.000)
Expected voxels per cluster, <k> = 81.441
Expected number of clusters, <c> = 7.74
FWEp: 7.510, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 16.3 16.2 14.7 mm mm mm; 10.9 10.8 9.8 {voxels}
Volume: 1685704 = 499468 voxels = 406.4 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1150.68 voxels)
Page 1



SPMresults: ./group/contrast_59
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.893	6	0.946	0.930	62	0.320	0.672	0.460	5.28	3.91	0.000	-8	-6	1
		0.946	0.930	62	0.320	0.754	0.460	5.11	3.83	0.000	-24	-60	-2
		0.995	0.930	21	0.575	0.997	0.982	4.09	3.30	0.000	-6	-30	43
		0.999	0.930	7	0.766	0.999	0.982	3.92	3.20	0.001	22	-25	43
		1.000	0.930	1	0.930	1.000	0.982	3.77	3.11	0.001	-24	-26	42
		1.000	0.930	2	0.891	1.000	0.982	3.75	3.10	0.001	46	-13	-30

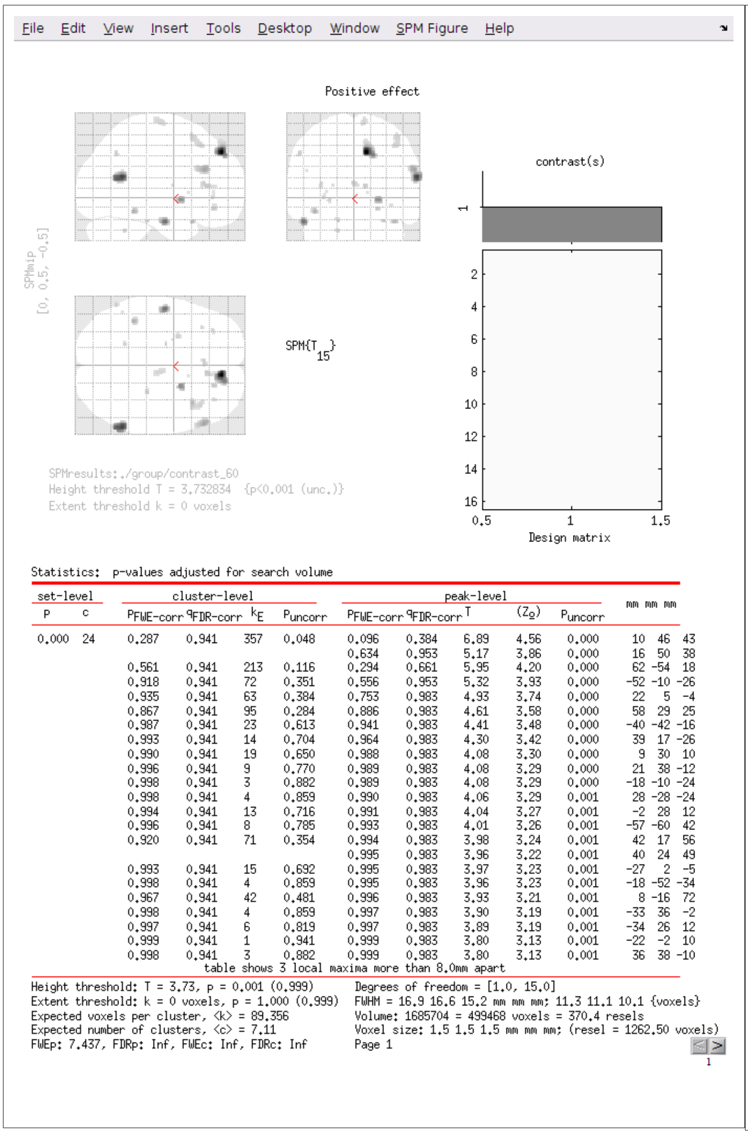
table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 1.000 (1.000)
Expected voxels per cluster, <k> = 67.796
Expected number of clusters, <c> = 9.14
FWEp: 7.654, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 15.4 15.0 14.0 mm mm mm; 10.3 10.0 9.3 {voxels}
Volume: 1685704 = 499468 voxels = 488.2 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 957.89 voxels)

Figure H.3. Syntactic effects: brain glass image and statistic tables for uncorrected p-value < 0.001 of pairs of syntactically different formulas over pairs of syntactically similar formulas at the left and the same contrast including interaction with structure over non structure of formulas on the right.

H.4) Branching effects: brain glass image and statistic tables for uncorrected p-value < 0.001 of pairs of left branched formulas over pairs of right branched formulas at the left, and pairs of formulas with a syntactic change from branched to top location of the unary operator over pairs of formulas with similar syntactic structure at the right.



SPMresults: ./group/contrast_60
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

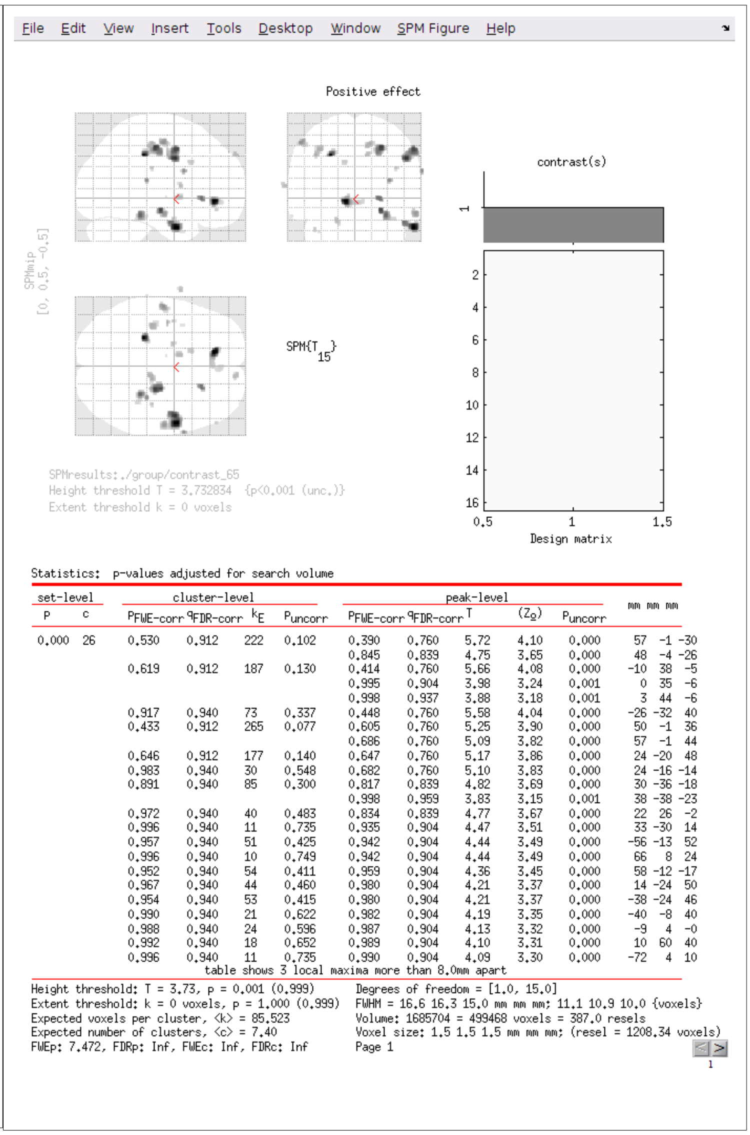
Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	24	0.287	0.941	357	0.048	0.096	0.384	6.89	4.56	0.000	10	46	43
						0.634	0.953	5.17	3.86	0.000	16	50	38
		0.561	0.941	213	0.116	0.294	0.661	5.95	4.20	0.000	62	-54	18
		0.918	0.941	72	0.351	0.556	0.953	5.32	3.93	0.000	-52	-10	-26
		0.935	0.941	63	0.384	0.753	0.983	4.93	3.74	0.000	22	5	-4
		0.867	0.941	95	0.284	0.886	0.983	4.61	3.58	0.000	58	29	25
		0.987	0.941	23	0.613	0.941	0.983	4.41	3.48	0.000	-40	-42	-16
		0.993	0.941	14	0.704	0.964	0.983	4.30	3.42	0.000	39	17	-26
		0.990	0.941	19	0.650	0.988	0.983	4.08	3.30	0.000	9	30	10
		0.996	0.941	9	0.770	0.989	0.983	4.08	3.29	0.000	21	38	-12
		0.998	0.941	3	0.882	0.989	0.983	4.08	3.29	0.000	-18	-10	-24
		0.998	0.941	4	0.859	0.990	0.983	4.06	3.29	0.001	28	-28	-24
		0.994	0.941	13	0.716	0.991	0.983	4.04	3.27	0.001	-2	28	12
		0.996	0.941	8	0.785	0.993	0.983	4.01	3.26	0.001	-57	-60	42
		0.920	0.941	71	0.354	0.994	0.983	3.98	3.24	0.001	42	17	56
						0.995	0.983	3.96	3.22	0.001	40	24	49
		0.993	0.941	15	0.692	0.995	0.983	3.97	3.23	0.001	-27	2	-5
		0.998	0.941	4	0.859	0.995	0.983	3.96	3.23	0.001	-18	-52	-34
		0.967	0.941	42	0.481	0.996	0.983	3.93	3.21	0.001	8	-16	72
		0.998	0.941	4	0.859	0.997	0.983	3.90	3.19	0.001	-33	36	-2
		0.997	0.941	6	0.819	0.997	0.983	3.89	3.19	0.001	-34	26	12
		0.999	0.941	1	0.941	0.999	0.983	3.80	3.13	0.001	-22	-2	10
		0.998	0.941	3	0.882	0.999	0.983	3.80	3.13	0.001	36	38	-10

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.73, p = 0.001 (0.999)
Extent threshold: k = 0 voxels, p = 1.000 (0.999)
Expected voxels per cluster, <k> = 89.356
Expected number of clusters, <c> = 7.11
FWEp: 7.437, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 16.9 16.6 15.2 mm mm mm; 11.3 11.1 10.1 {voxels}
Volume: 1685704 = 499468 voxels = 370.4 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1262.50 voxels)
Page 1



SPMresults: ./group/contrast_65
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.000	26	0.530	0.912	222	0.102	0.390	0.760	5.72	4.10	0.000	57	-1	-30
						0.845	0.839	4.75	3.65	0.000	48	-4	-26
		0.619	0.912	187	0.130	0.414	0.760	5.66	4.08	0.000	-10	38	-5
						0.995	0.904	3.98	3.24	0.001	0	35	-6
						0.998	0.937	3.88	3.18	0.001	3	44	-6
		0.917	0.940	73	0.337	0.448	0.760	5.58	4.04	0.000	-26	-32	40
		0.433	0.912	265	0.077	0.605	0.760	5.25	3.90	0.000	50	-1	36
						0.686	0.760	5.09	3.82	0.000	57	-1	44
		0.646	0.912	177	0.140	0.647	0.760	5.17	3.86	0.000	24	-20	48
		0.983	0.940	30	0.548	0.682	0.760	5.10	3.83	0.000	24	-16	-14
		0.891	0.940	85	0.300	0.817	0.839	4.82	3.69	0.000	30	-36	-18
						0.998	0.959	3.83	3.15	0.001	38	-38	-23
		0.972	0.940	40	0.483	0.834	0.839	4.77	3.67	0.000	22	26	-2
		0.996	0.940	11	0.735	0.935	0.904	4.47	3.51	0.000	33	-30	14
		0.957	0.940	51	0.425	0.942	0.904	4.44	3.49	0.000	-56	-13	52
		0.996	0.940	10	0.749	0.942	0.904	4.44	3.49	0.000	66	8	24
		0.952	0.940	54	0.411	0.959	0.904	4.36	3.45	0.000	58	-12	-17
		0.967	0.940	44	0.460	0.980	0.904	4.21	3.37	0.000	14	-24	50
		0.954	0.940	53	0.415	0.980	0.904	4.21	3.37	0.000	-38	-24	46
		0.990	0.940	21	0.622	0.982	0.904	4.19	3.35	0.000	-40	-8	40
		0.988	0.940	24	0.596	0.987	0.904	4.13	3.32	0.000	-9	4	-0
		0.992	0.940	18	0.652	0.989	0.904	4.10	3.31	0.000	10	60	40
		0.996	0.940	11	0.735	0.990	0.904	4.09	3.30	0.000	-72	4	10

table shows 3 local maxima more than 8.0mm apart

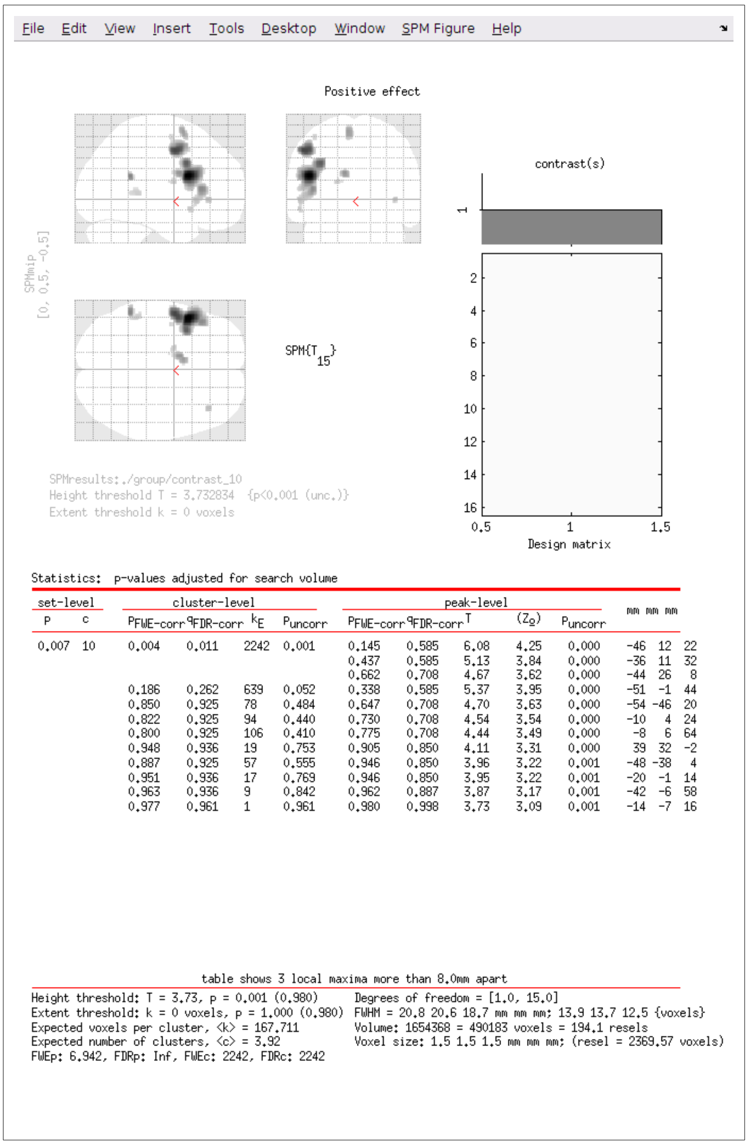
Height threshold: T = 3.73, p = 0.001 (0.999)
Extent threshold: k = 0 voxels, p = 1.000 (0.999)
Expected voxels per cluster, <k> = 85.523
Expected number of clusters, <c> = 7.40
FWEp: 7.472, FDRp: Inf, FWEc: Inf, FDRc: Inf

Degrees of freedom = [1.0, 15.0]
FWHM = 16.6 16.3 15.0 mm mm mm; 11.1 10.9 10.0 {voxels}
Volume: 1685704 = 499468 voxels = 387.0 resels
Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1208.34 voxels)
Page 1

Figure H.4. Branching effects: brain glass image and statistic tables for uncorrected p-value < 0.001 of pairs of left branched formulas over pairs of right branched formulas at the left, and pairs of formulas with a syntactic change from branched to top location of the unary operator over pairs of formulas with similar syntactic structure at the right.

Appendix I. Simple Composition Brain Activations

I.1) Language network activation: brain glass image and statistic tables for uncorrected p-value < 0.001 of complete phrases over one word trials.



SPMresults: ./group/contrast_10
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

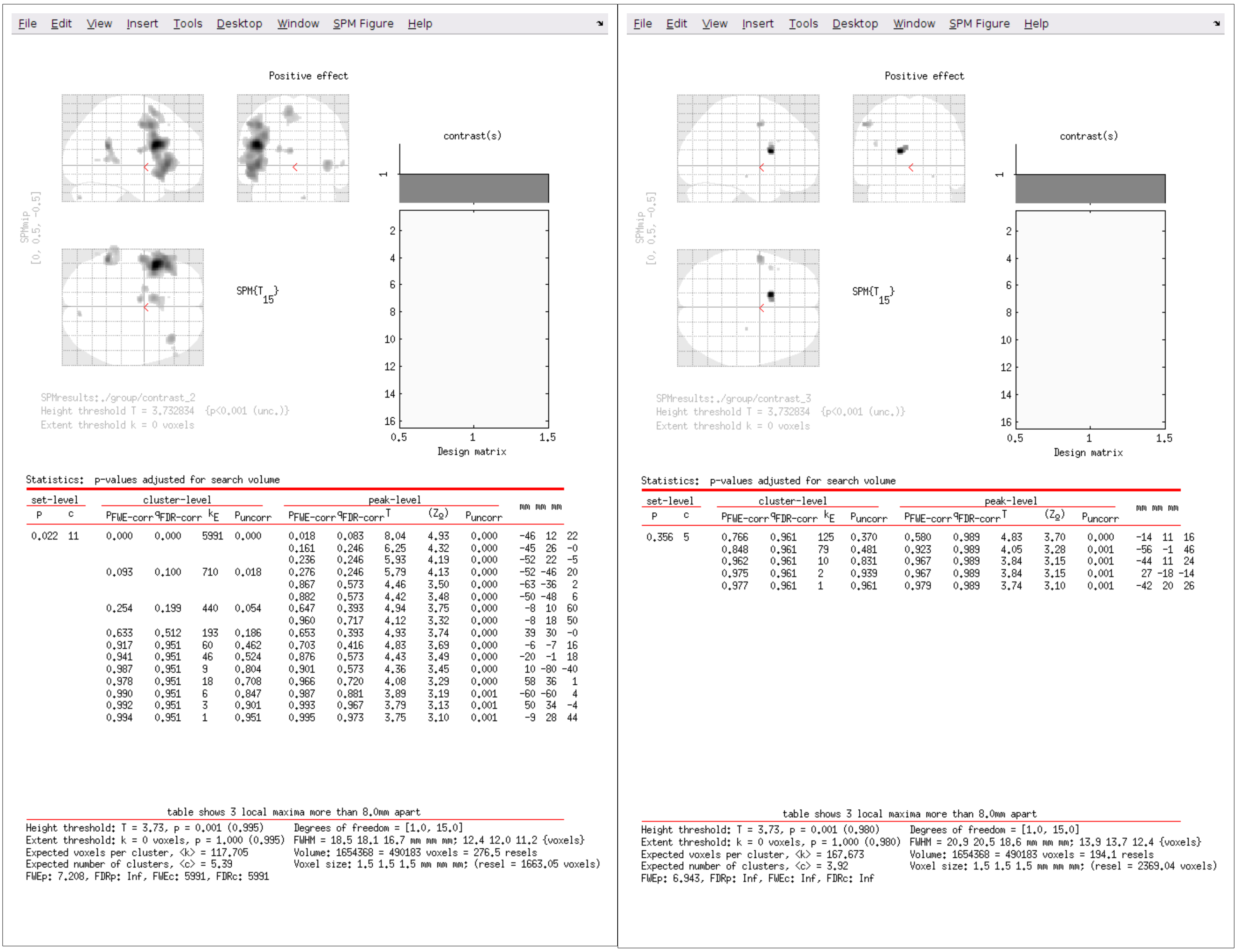
set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.007	10	0.004	0.011	2242	0.001	0.145	0.585	6.08	4.25	0.000	-46	12	22
						0.437	0.585	5.13	3.84	0.000	-36	11	32
						0.662	0.708	4.67	3.62	0.000	-44	26	8
		0.186	0.262	639	0.052	0.338	0.585	5.37	3.95	0.000	-51	-1	44
		0.850	0.925	78	0.484	0.647	0.708	4.70	3.63	0.000	-54	-46	20
		0.822	0.925	94	0.440	0.730	0.708	4.54	3.54	0.000	-10	4	24
		0.800	0.925	106	0.410	0.775	0.708	4.44	3.49	0.000	-8	6	64
		0.948	0.936	19	0.753	0.905	0.850	4.11	3.31	0.000	39	32	-2
		0.887	0.925	57	0.555	0.946	0.850	3.96	3.22	0.001	-48	-38	4
		0.951	0.936	17	0.769	0.946	0.850	3.95	3.22	0.001	-20	-1	14
		0.963	0.936	9	0.842	0.962	0.887	3.87	3.17	0.001	-42	-6	58
		0.977	0.961	1	0.961	0.980	0.998	3.73	3.09	0.001	-14	-7	16

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.980)          Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.980)      FWHM = 20.8 20.6 18.7 mm mm mm; 13.9 13.7 12.5 {voxels}
Expected voxels per cluster, <k> = 167.711             Volume: 1654368 = 490183 voxels = 194.1 resels
Expected number of clusters, <c> = 3.92                Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 2369.57 voxels)
FWEp: 6.942, FDRp: Inf, FWEc: 2242, FDRc: 2242
```

Figure I.1. Language network activation: brain glass image and statistic tables for uncorrected p-value < 0.001 of complete phrases over one word trials.

I.2) Increasing complexity: brain glass image and statistic tables for uncorrected p-value < 0.001 of linear phrase complexity at the left, and linear list complexity at the right.



SPMresults: ./group/contrast_2
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level					mm	mm	mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}			
0.022	11	0.000	0.000	5991	0.000	0.018	0.083	8.04	4.93	0.000	-46	12	22
						0.161	0.246	6.25	4.32	0.000	-45	26	-0
						0.236	0.246	5.93	4.19	0.000	-52	22	-5
		0.093	0.100	710	0.018	0.276	0.246	5.79	4.13	0.000	-52	-46	20
						0.867	0.573	4.46	3.50	0.000	-63	-36	2
						0.882	0.573	4.42	3.48	0.000	-50	-48	6
		0.254	0.199	440	0.054	0.647	0.393	4.94	3.75	0.000	-8	10	60
						0.960	0.717	4.12	3.32	0.000	-8	18	50
		0.633	0.512	193	0.186	0.653	0.393	4.93	3.74	0.000	39	30	-0
		0.917	0.951	60	0.462	0.703	0.416	4.83	3.69	0.000	-6	-7	16
		0.941	0.951	46	0.524	0.876	0.573	4.43	3.49	0.000	-20	-1	18
		0.987	0.951	9	0.804	0.901	0.573	4.36	3.45	0.000	10	-80	-40
		0.978	0.951	18	0.708	0.966	0.720	4.08	3.29	0.000	58	36	1
		0.990	0.951	6	0.847	0.987	0.881	3.89	3.19	0.001	-60	-60	4
		0.992	0.951	3	0.901	0.993	0.967	3.79	3.13	0.001	50	34	-4
		0.994	0.951	1	0.951	0.995	0.973	3.75	3.10	0.001	-9	28	44

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.995)        Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.995)    FWHM = 18.5 18.1 16.7 mm mm mm; 12.4 12.0 11.2 {voxels}
Expected voxels per cluster, <k> = 117.705           Volume: 1654368 = 490183 voxels = 276.5 resels
Expected number of clusters, <c> = 5.39              Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1663.05 voxels)
FWEp: 7.208, FDRp: Inf, FWEc: 5991, FDRc: 5991
```

SPMresults: ./group/contrast_3
Height threshold T = 3.732834 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Statistics: p-values adjusted for search volume

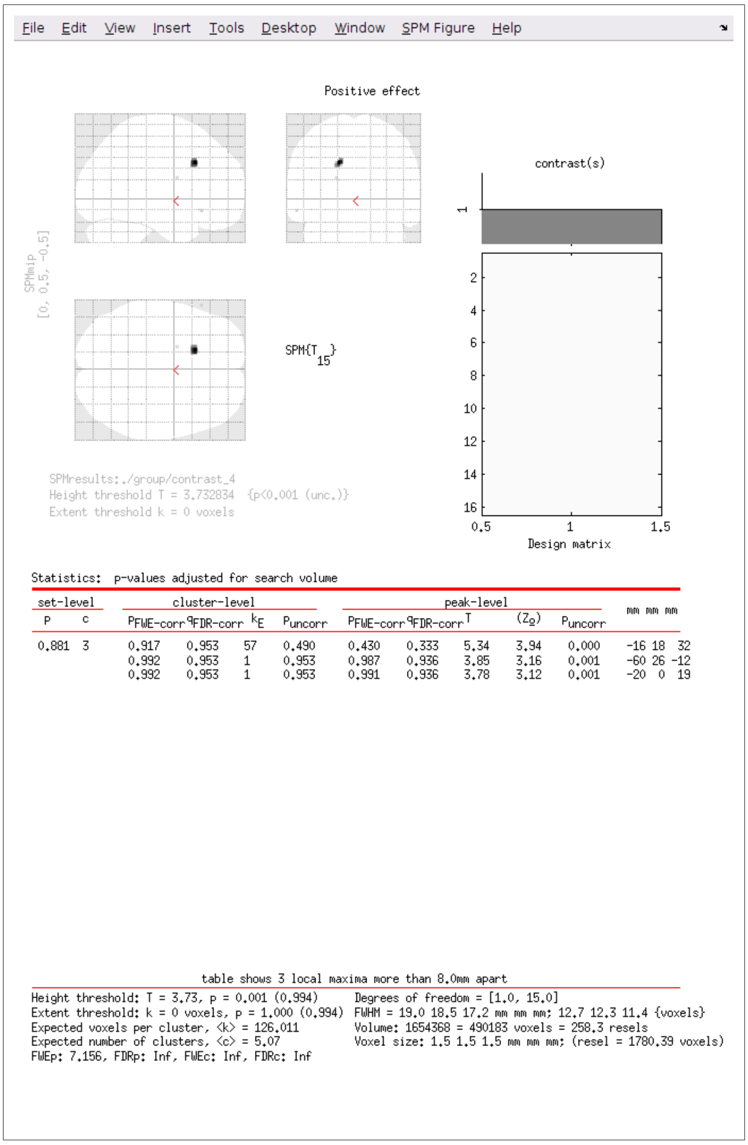
set-level		cluster-level				peak-level					mm	mm	mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}			
0.356	5	0.766	0.961	125	0.370	0.580	0.989	4.83	3.70	0.000	-14	11	16
		0.848	0.961	79	0.481	0.923	0.989	4.05	3.28	0.001	-56	-1	46
		0.962	0.961	10	0.831	0.967	0.989	3.84	3.15	0.001	-44	11	24
		0.975	0.961	2	0.939	0.967	0.989	3.84	3.15	0.001	27	-18	-14
		0.977	0.961	1	0.961	0.979	0.989	3.74	3.10	0.001	-42	20	26

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.980)        Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.980)    FWHM = 20.9 20.5 18.6 mm mm mm; 13.9 13.7 12.4 {voxels}
Expected voxels per cluster, <k> = 167.673           Volume: 1654368 = 490183 voxels = 194.1 resels
Expected number of clusters, <c> = 3.92              Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 2369.04 voxels)
FWEp: 6.943, FDRp: Inf, FWEc: Inf, FDRc: Inf
```

Figure I.2. Increasing complexity: brain glass image and statistic tables for uncorrected p-value < 0.001 of linear phrase complexity at the left, and linear list complexity at the right.

I.3) Structure effect: brain glass image and statistic tables for uncorrected p-value < 0.001 of linear phrase complexity over linear list complexity.



```
SPMresults: ./group/contrast_4
Height threshold T = 3.732834  {p<0.001 (unc.)}
Extent threshold k = 0 voxels
```

Statistics: p-values adjusted for search volume

set-level		cluster-level				peak-level							
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_{\equiv})$	p_{uncorr}	mm	mm	mm
0.881	3	0.917	0.953	57	0.490	0.430	0.333	5.34	3.94	0.000	-16	18	32
		0.992	0.953	1	0.953	0.987	0.936	3.85	3.16	0.001	-60	26	-12
		0.992	0.953	1	0.953	0.991	0.936	3.78	3.12	0.001	-20	0	19

table shows 3 local maxima more than 8.0mm apart

```
Height threshold: T = 3.73, p = 0.001 (0.994)        Degrees of freedom = [1.0, 15.0]
Extent threshold: k = 0 voxels, p = 1.000 (0.994)    FWHM = 19.0 18.5 17.2 mm mm mm; 12.7 12.3 11.4 {voxels}
Expected voxels per cluster, <k> = 126.011           Volume: 1654368 = 490183 voxels = 258.3 resels
Expected number of clusters, <c> = 5.07              Voxel size: 1.5 1.5 1.5 mm mm mm; (resel = 1780.39 voxels)
FWEp: 7.156, FDRp: Inf, FWEc: Inf, FDRc: Inf
```

Figure I.3. Structure effect: brain glass image and statistic tables for uncorrected p-value < 0.001 of linear phrase complexity over linear list complexity.