4

Results

4.1

Systematic Search

The systematic search in the neurosynth.org database retrieved 476 articles on working memory. Among these, 28 assessed phonological or auditory working memory. Table 2 depicts the neurosynth.org database results.

Table 2. Total of studies retrieved from theneurosynth.org database, divided by type of study based on the inclusion and exclusion criteria.

Variable	Number of articles
Visual working memory	39
Studies that usedregions of interest	16
Unclear behavioral methods	13
Studies that did not assessworking memory directly/clinical samples	380
Phonological, auditory, or verbal working memory	28
TOTAL	476

When other databases were considered, we found 233 articles, but 137 of them were already in the neurosynth.org database. Considering the other 96 articles, two assessed phonological working memory. The final result of the systematic search yielded 30 articles that studied phonological working memory. Table 3 shows the results of the other databases.

Table 3. Total of retrieved studies from other databases other than neurosynth.org divided by database.

Database	Number of articles
PLoS One	2
Oxford	9
SpringerLink	54
Elsevier	0
Cambridge	0
Wiley& Sons	31
Total	96

The total number of scans (i.e., number of participants) was 596. The minimum sample size was eight participants (Stoodley, Valera, & Schmahmann, 2012). The maximum sample size was 87 participants (Spielberg et al., 2011), with an average of 19.9 participants per study (standard deviation = 16.5). With regard to the fMRI machine that was used to collect the data, 11 used 1.5 T scanners, 17 used 3.0 T scanners, and two used 4.0 T scanners. With regard to the time of publication of the phonological working memory studies, the oldest study was published in 2000 (Martinkauppi, Rama, Aronen, Korvenoja, & Carlson, 2000), and the newest study was published in 2013 (Dima, Jogia, & Frangou, 2013). Table 4 presents the results of the articles retrieved based on the established criteria.

Table 4. Authorship, year of publication, sample size, machine used for neuroimaging, task description, and contrast observed in the experimental protocol.

Author	Year	Sample size	Machine	Task description	Contrast
Barry JG, Sabisch B, Friederici AD, Brauer J	2011	16	3.0 T	Participants were asked to engage in two tasks. The first was nonword repetition with 19 pseudowords that varied from one to five syllables. The second task was a complex delayed match-to-sample. The participant first both heard and saw a pair of words for 1 s, and then another pair of words was presented for 1 s(encoding phase). A rehearsal phase followed for 4, 6, or 8 s. The rehearsed stimuli could be the words that were heard(an ear appeared on the screen), the words that were read (an eye appeared on the screen), or neither (a hand appeared on the screen). Finally, the participant had to respond whether the targets were both correct (the same of the two stimuli), one correct and one wrong, or both wrong.	Main effect: correlation between nonword repetition and encoding, nonword repetition, and recognition Interaction: correlation between nonword repetition and encoding minus baseline, nonword repetition and recognition minus baseline, nonword repetition and encoding minus recognition

Author	Year	Sample	Machine	Task Description	Contrast
Bunge SA, Ochsner KN, Desmond JE, Glover GH, Gabrieli JD	2001	16	3.0 T	The Sternberg Item Recognition paradigm is a computerized protocol that consists of presenting a single stimulus or set of stimuli at one time. After a short period of maintenance, participants must respond whether a probe item was among the previously presented stimuli. Sets in this experiment have one, four, or six stimuli in each trial (Load 1, 4, or 6). Stimuli from the previous two trials were not presented as either stimuli or probes. However, in this experiment, one condition was established using probes from the immediately prior trial. This conditions was called Load 4 High Recency.	Main effect: Load 6 minus Load 4;Load 4 High Recency minus Load 4 Interaction: all load-related activation
Chein JM, Fiez JA	2001	12	1.5 T	The task consisted of a list of five words (for each type of stimulus) that was serially presented for 8 s(encoding phase). A 20 s maintenance phase followed. A retrieval task was then conducted, in which a probe word was presented, and the participant had to answer "yes" or "no" regarding whether the probe was presented previously during encoding. The number of syllables, distinction or similarity between phonetics, and word *vs.* nonword conditions were only treated in terms of difficulty rather than analyzed one at a time. The intertrial interval was 20 s(baseline).	Main effect: encoding minus baseline; maintenance minus baseline; retrieval minus baseline; encoding plus maintenance minus baseline; encoding plus retrieval minus baseline; maintenance plus retrieval minus baseline; encoding plus maintenance plus retrieval minus baseline Interaction: encoding minus maintenance; encoding minus retrieval; maintenance minus retrieval; maintenance in the four difficulty conditions
Dima D, Jogia J, Frangou S	2013	40	1.5 T	Participants were asked to respond to 0-, 1-, 2-, and 3-back procedures for letters in a total of 252 trials.	Main effect: 1-back minus 0-back; 2-back minus 0-back; 3-back minus 0-back
Fleck MS, Daselaar SM, Dobbins IG, Cabeza R	2006	14	4.0 T	Participants were asked to memorize a list of words before scanning. The word recognition task consisted of deciding whether a word was seen in the list. The visual perception task consisted of deciding whether one color or another was predominant. After each trial, the participants were asked to evaluate their confidence in the responses they gave.	Main effect: four confidence levels in both tasks Interaction: recognition minus visual perception in the four confidence levels

Table 4. Continued.

Author	Year	Sample	Machine	Task description	Contrast
Gluber	2001	11	3.0 T	Four letters were simultaneously presented as stimuli for 1 s. A 4 s maintenance phase followed. The participant then had to respond whether the probe letter was present in the previously presented stimuli. The intertrial interval was 1 s(baseline). Two conditions were implemented: in the first condition, the participants were asked to suppress any articulatory rehearsal during the maintenance phase; in the second condition, the participants were asked to articulate the string of letters during maintenance.	Main effect: articulatory suppression minus rest; non-articulatory suppression minus rest Interaction: articulatory suppression plus non-articulatory suppression minus rest
Hester R, D'Esposito M, Cole MW, Garavan H	2007	13	4.0 T	Participants were asked to pay attention to a list of five letters on a screen (encoding phase). After this phase, a maintenance phase (8-12s) occurred. A probe item (one letter among the ones that were presented in the list) was shown for 2 s (selection phase), and a 6-10 s preparation phase was followed by a stable recall or variable recall task. In the stable recall task, the participant had to retrieve the item that followed the probe in the list using a simple serial code for the response buttons. In the variable condition, the participants had to respond using randomly assigned buttons according to the order that was presented on the screen.	Main effect: encoding minus rest; preparation plus selection minus rest; maintenance minus rest Interaction: (variable minus stable) minus rest
Karlsgodt KH, Shirinyan D, van Erp TG, Cohen MS, Cannon TD	2005	13	3.0 T	Participants were asked to pay attention to a list of five words that were presented one-by-one on the screen for 1 s each (encoding phase). After an 8-sdelay (maintenance phase), the participants were asked to match 8-10 probe words to the list that was previously shown (retrieval) by pressing "yes" or "no" buttons.	Main effect: encoding minus rest; maintenance minus rest; retrieval minus rest Interaction: encodingminus maintenance minus retrieval
Leung AW, Alain C	2011	16	3.0 T	Sounds were presented in three location conditions (-90º, 0º, 90º). Four experimental conditions were extracted using a 2×2 design: (1) category *vs*. location and (2) n-back 1 *vs*. n-back 2.	Main effect: Category minus Location Interaction: Category2-back minus Category1-back > Location2-back minus Location1-back
Ma L, Steinberg JL, Hasan KM, Narayana PA, Kramer LA, Moeller FG	2012	18	3.0 T	An Immediate Memory Task (IMT) consisted of a matching-to-sample task in which a target was presented and matching stimuli were presented after the target. A Delayed Memory Task (DMT) consisted of the same procedure, but a distraction (string 000000) was presented three times between the target and stimuli.	Main effect: DMT minusIMT-calledDI Interaction: DI7minusDI5; DI3minusDI5

Table 4. Continued.

Author	Year	Sample	Machine	Task description	Contrast
Majerus S, Van der Linden M, Collette F, Laureys S, Poncelet M, Degueldre C, Delfiore G, Luxen A, Salmon E	2005	12	3.0 T	Participants were asked to repeat words and nonwords that were randomly presented in three different blocks, with one 20-trial block for each condition. After the first section of tasks, the participants were exposed to another 20 stimuli six times. This phase was called familiarization. Those stimuli were different from the first 20 stimuli. After familiarization, the participants performed the task again with the familiar stimuli.	Main effect: Familiarization minus Non-familiarization for words and nonwords; Non-familiarizationminus Familiarization for words and nonwords
Martinkauppi S, Rama P, Aronen HJ, Korvenoja A, Carlson S	2000	10	1.5 T	Sounds were presented in three location conditions (left, right, and middle).The participants had to perform n-back tasks for the location of the sounds. A control experiment consisted of a visual n-back of 2, in which a white square was presented in three possible locations on a black screen (left, top-middle, and right), and the participants had to respond 2-back for the locations.	Main effect: 3-back minus 1-back; 2-back minus 1-back Interaction: Auditory 2-back minus Visual 2-back
Marvel CL, Desmond JE	2010	16	3.0 T	In a simple matching-to-sample condition (called match condition), sets of two or six letters were presented to the participants for 2 s (encoding phase). A 4-6 s interstimulus interval then occurred (maintenance phase). One single letter was then presented as the target for 1 s, and the participants had to judge whether this letter belonged to the set of letters that was previously presented (retrieval phase). In the complex matching-to-sample condition (called executive condition), the same procedure was adopted with regard to the target being a letter that was two positions before in the alphabet(e.g., if the letter *f* was shown in a set, and the letter *d* appeared as the target, then the participant has to press the "yes" button because *d* is two positions before *f* in the alphabet).	Main effect: executive condition minus match condition Interaction: executive condition minus match condition during encoding phase; executive condition minus match condition during maintenance phase; executive condition minus match condition during retrieval phase
Marvel CL, Desmond JE	2012	16	3.0 T	Two conditions of a verbal version of the Sternberg working memory task were designed. In the first condition, one or two strings of three elements of the same letter (e.g., F-F-F and Q-Q-Q) were presented for 1 s(encoding phase). A maintenance phase then followed, varying from 4 to 6 s. A retrieval task was then conducted, asking the participants to judge whether a probe letter was presented in the previously shown string. The second condition was identical to the first condition, with the exception that the probe letter would have to be two positions ahead in the alphabet of the presented letter instead of the encoded letter (e.g., if the stimuli were F-F-F, then the correct probe letter should be H because it is two positions ahead in the alphabet).	Interaction: 2-target minus 1-target in the second condition (manipulation) minus 2-target minus 1-target in the first condition (storage)

Table 4. Continued.

Author	Year	Sample	Machine	Task description	Contrast
McNab F, Leroux G, Strand F, Thorell L, Bergman S, Klingberg T	2008	11	1.5 T	Verbal DMT. The participants were required to encode five letters that were sequentially presented for 500 ms each. A 1 s maintenance phase followed. A probe cue that consisted of a number and a letter was then shown, and the participant had to decide whether the serial position of the letter in the probe cue was the same compared with the serial position of the previously presented stimulus (e.g., for the sequential stimuli G, D, W, F, and M, the probe cue 3:W would be a correct match, as well as 2:D, 4:F, and so on). A control task consisted of always presenting the letter *a* in lower case, and the letter *A* in upper case was the cue probe.	Main effect: DMT minus control; Conjoint Visual and Verbal DMT plus Go/No-Go Task; Conjoint Visual and Verbal DMT plus Flanker Task; Conjoint Verbal and Visual DMT plus Go/No-Go Task plus Flanker Task
Nee DE, Jonides J	2011	25	3.0 T	The task consisted of a DMT using a list of six four-letter words for each trial. Words were serially presented (500 ms per word), followed by a 300 ms mask (%%%%) and a retrieval task with a probe word. The probe word was either in activated long-term memory (a LTM; first presented stimuli in the list) at the region of direct access (RDA; positions 5 to 2) or at the focus of attention (FA; the last presented word), according to Oberauer's (2002) 3-state model of memory. The intertrial interval was 4-7 s(baseline).	Main effect: aLTM minus baseline; RDA minus baseline; FA minus baseline Interaction: FA minus RDA; FA minus aLTM, RDA minus aLTM
Newton AT, Morgan VL, Rogers BP, Gore JC	2011	8	3.0 T	Participants were asked to respond to 0-, 1-, 2-, and 3-back procedures for letters in 20 trials in each condition.	Interaction: 3-back minus 2-back minus 1-back minus 0-back minus rest
Novais-Santos S, Gee J, Shah M, Troiani V, Work M, Grossman M	2007	19	3.0 T	Participants were asked to read Direct Object and SC sentences in a "moving window" paradigm. Phrases from the sentences were presented in the following order: initial-phrase; verb-phrase; noun-phrase (working memory[WM]-phrase/50% of the cases); concluding-phrase. In 20% of the trials, the participants were asked to answer simple comprehension questions to ensure they were paying attention. However, this part of the procedure was not entered into the analysis. Sentences with a WM phrase were considered "more WM" because they demanded higher loads to be comprehended.	Main effect: SC minus DO; More WM minus Less WM
Relander K, Rama P	2009	10	3.0 T	A matching-to-sample task was conducted using auditory input. The participants were asked to respond whether the target matched the stimuli in three different situations: only voices, only words, and control. A word that consisted of "Voice,""Word," or "Control" was shown on the screen 3 s before each trial, so participants could know whether they should answer based on the speaker's voice, the spoken word, or no response (control condition) in each particular trial. The stimuli then appeared for 0.9-1.2 s (encoding phase), followed by an interstimulus interval of 4.5 s(maintenance phase). The target was then presented, and the participants had 3.0 s to answer (retrieval phase).	Main effect: voice minus control; words minus control Interaction: voice minus word during encoding; voice minus word during maintenance; voice minusword during retrieval

Table 4. Continued.

Author	Year	Sample	Machine	Task description	Contrast
Rodriguez-Jimenez R, Avila C, Garcia-Navarro C, Bagney A, Aragon AM, Ventura-Campos N, Martinez-Gras I, Forn C, Ponce G, Rubio G, Jimenez-Arriero MA, Palomo T	2009	13	1.5 T	Participants were asked to respond to 2-back procedures for letters that were presented in two different conditions (visual and verbal)in 23 trials in each condition.	Main effect: Auditory plus Visual minus rest Interaction: Auditory minus Visual
Rose M, Schmid C, Winzen A, Sommer T, Buchel C	2005	14	1.5 T	All stimuli consisted of a scene that was superposed by a letter. Different levels of visibility (0%, 25%, 50%, 75%, 100%) of the scenes were covered by abstract patterns and with color reduced by RGB color palette were used during both n-back tasks. The participants were asked to ignore the background for task purposes and only respond to the n-back procedure.	Main effect: 2-back minus 1-back; random effects for the 5 levels of visibility Interaction: random effects for visibility during 2-back minus 1-back
Rudner M, Fransson P, Ingvar M, Nyberg L, Ronnberg J	2007	13	1.5 T	Participants performed a 2-back task for each condition: sign language, speech language, and both (binding).	Main effect: sign language minus rest; speech minus rest; binding minus rest Interaction: sign minus speech; binding minus (sign plus speech)
Sabb FW, Bilder RM, Chou M, Bookheimer SY	2007	17	3.0 T	In each trial, the participants saw a list of words with 3 (low load), 5 (medium load), or 7 (high load) items that were presented one at a time in the middle of the screen. The participants had to respond as quickly as possible if each list item was a living or nonliving object. After presenting the list, the participants were asked a probe question regarding whether a particular item in the set was a member of a particular category (e.g., “Was the second item a fruit?”). Some items were used as priming for others in a hierarchical level (e.g., “guitar” and “violin”).	Main effect: low load minus rest; medium load minus rest; high load minus rest; primed items minus unprimed items in low load; primed items minus unprimed items in medium load; primed items minus unprimed items in high load Interaction: primed minus unprimed in all loads
Schaefer A, Braver TS, Reynolds JR, Burgess GC, Yarkoni T, Gray JR	2006	53	1.5 T	The participants first watched a film to increase emotional arousal or maintain it (control condition). The participants were then asked to respond in 1- and 3-back procedures for words (concrete nouns) and faces in 21 trials in each condition.	Main effect: 3-back minus1-back Interaction: 3-back minus 1-back in either emotional or control condition

Table 4. Continued.

Author	Year	Sample	Machine	Task description	Contrast
Spielberg JM, Miller GA, Engels AS, Herrington JD, Sutton BP, Banich MT, Heller W	2011	87	3.0 T	Participants answered three questionnaires to determine whether they presented personality traits that were closer to an approach temperament or avoidance temperament. A color-word Stroop test with 256 trials was then conducted to determine how different levels of temperament traits would be reflected in task performance.	Main effect: congruent words minus incongruent words; congruent words minus incongruent words in approach temperament; congruent minus incongruent words in avoidance temperament Interaction: congruent words minus incongruent words in avoidance minus approach temperaments
Stoodley CJ, Valera EM, Schmahmann JD	2012	8	3.0 T	Participants were asked to respond in 0- and 2-back procedures for letters in 16 trials in each condition.	Main effect: 2-back minus 0-back
Suchan B, Linnewerth B, Koster O, Daum I, Schmid G	2006	13	1.5 T	Participants were asked to respond in 0- and 2-back procedures for concrete nouns or pictures of these nouns that were presented in four different conditions: visual-visual (VV; only pictures were presented), auditory-auditory (AA; only words were spoken), auditory-visual (AV; the noun was first spoken and the target was the picture associated with this noun), and visual-auditory (VA; the opposite of the AV condition). Each condition had 30 trials.	Main effect: VV minus AA Interaction: VA minus AA
Vandewalle G, Gais S, Schabus M, Balteau E, Carrier J, Darsaud A, Sterpenich V, Albouy G, Dijk DJ, Maquet P	2007	18	3.0 T	An auditory 2-back design was established using nine phonologically different monosyllabic consonants. The experimental condition was the exposure of one eye of the participant to a green or blue light for 18 min before starting the 2-back experiment on two different days and in three different sessions during each day. The participants were randomly assigned to green or blue light on the first day.	Main effect: blue light exposure minus green light exposure Interaction: blue light minus green light in session 1 minus session 2; blue light minus green light in session 2 minus session 3
Woodward TS, Cairo TA, Ruff CC, Takane Y, Hunter MA, Ngan ET	2006	18	1.5 T	Strings of 2, 4, 6, or 8 letters were presented to the participants for 4 s(encoding phase). A maintenance phase then followed, varying from 3 to 5 s. The retrieval task consisted of answering “yes” or “no” regarding whether a target probe was present in the string of letters. The intertrial interval was 3 s.	Main effect: encoding minus rest, maintenance minus rest Interaction: encoding minus maintenance

Table 4. Continued.

4.2

ALE meta-analysis results

The results of the ALE meta-analysis considered the average activation (and average errors) in 30 phonological working memory studies (*n* =596). For the analyses, we categorized three types of stimuli: (1) tones (nonhuman sounds), (2) syllables (letters or phonemes without meaning),and (3) words (meaningful morphemes, nonwords, and phrases). Figure 6 shows activation related to tones (blue is peripheral and green is the center of activation), words (orange is peripheral and white is the center of activation), and syllables (red is peripheral and white is the center of activation) as the types of stimuli. Overlapping regions are depicted in yellow-red colors. Figure 7 depicts only overlapping regions of words>syllables.

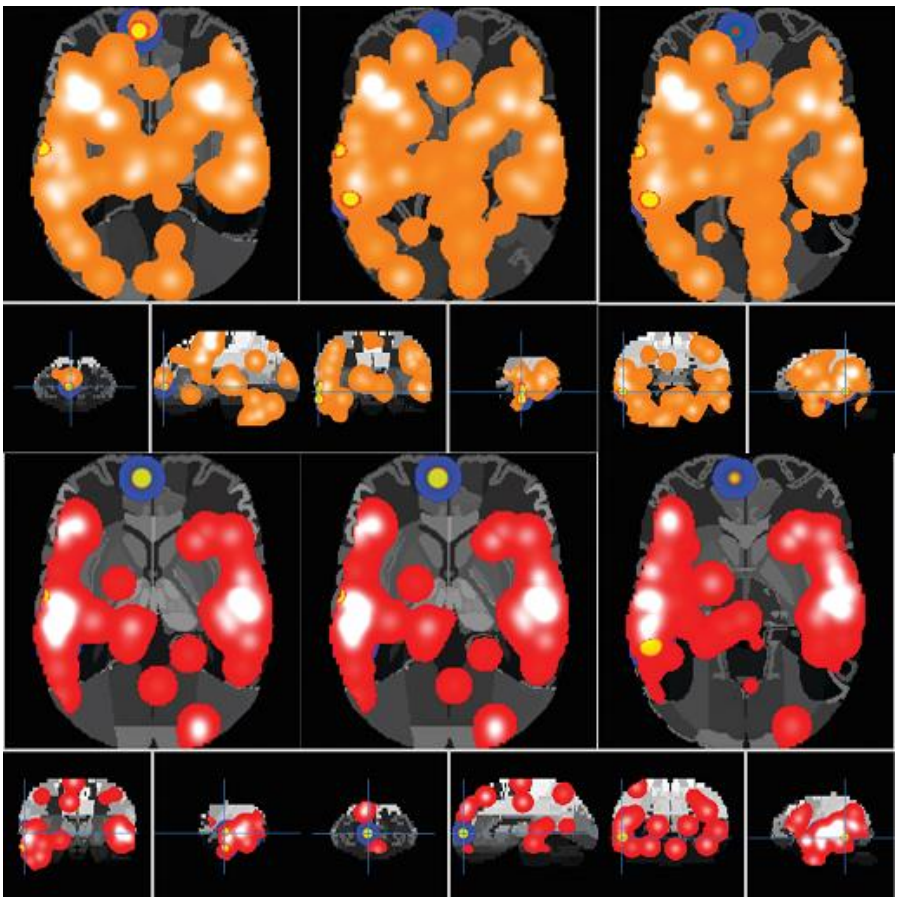

Fig. 6. Average activation of the participants (*n*=596) and standard deviation represented by types of stimuli.

The results in Figure6 show that important regions that are associated with working memory overlap while executing experimental tasks. With regard to tones and syllables, significant activation of the left prefrontal cortex is seen, going from the anterior to dorsolateral portions. Another notable aspect is the overlapping activation of both the auditory association cortex at the temporal lobe and supramarginal gyrus (part of Wernicke's area). Another interesting aspect is significant activation in parietal and premotor areas but only when the participants were executing tasks using syllables, which can explain why no overlap was observed in these regions between types of stimuli. The same phenomenon occurred when tones and words that overlapped but at a higher intensity. Almost the whole brain was activated in working memory tasks that used words or phrases as stimuli, but the overlapping regions were restricted to what was found with syllables>tones, with higher overlapping of the auditory association cortex. These results are discussed in more detail below.

Figure 7 shows only the overlapping activation of words>syllables. The results suggest double activation of the prefrontal cortex in both the anterior and dorsolateral portions of this region. They also appear to be separate, which likely is associated with the type of analyses that were conducted (i.e., using clusters to find centers of activation). Additionally, the bilateral medial frontal cortex was activated during tasks with words and syllables. Other regions of activation included the supramarginal gyrus (part of Wernicke's area), a small portion of the anterior intraparietal sulcus, and bilateral activation of the cingulate cortex (anterior insula). The pars opecularis (part of Broca's area) was also activated in working memory tasks using both types of stimuli.

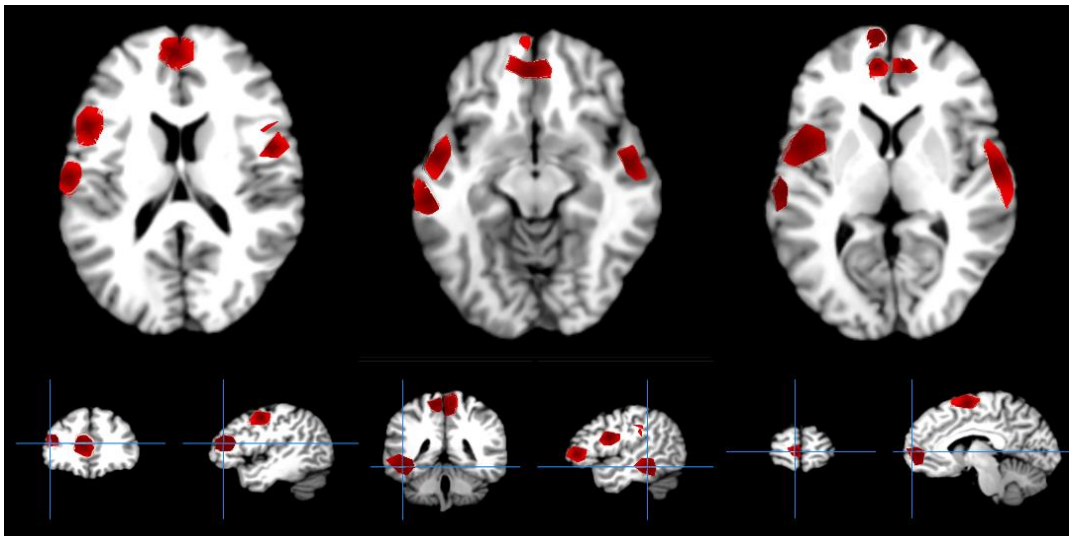

Fig. 7. Significantly overlapping regions using the contrastwords>syllables,excluding tones. The centers of activation were situated in Brodmann areas 6, 10, 40, 42, 44, and 46.

Using the ALE cluster method of contrast, we subtracted from the most widespread activation (involving words); the other activation likelihoods were used to generate the final results that are depicted in Figure 8. Contrast is represented by words>syllables>tones (p<0.001). The centers of activation were the left anterior prefrontal cortex (Brodmann area 10), left fusiform gyrus at the surface of the temporal lobe (Brodmann area 37), and left anterior transverse temporal gyrus at the surface of the temporal lobe (Brodmann area 42). These results are discussed in more detail below with regard to the function of these regions and the theoretical model they may support.

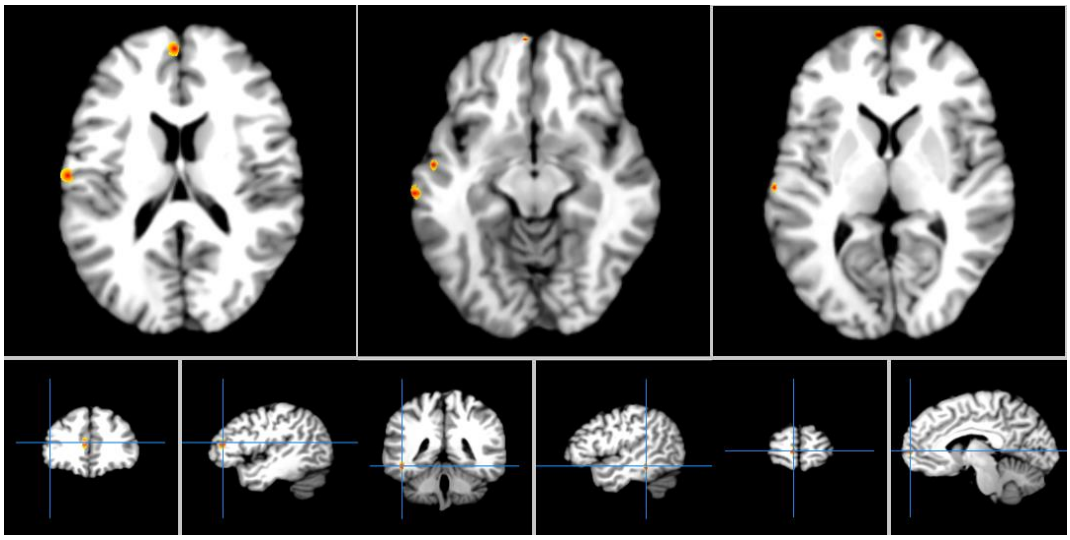

Fig. 8. Significantly overlapping regions using the contrast words>syllables>tones. The centers of activation were situated in Brodmann areas 10, 37, and 42.