

TITLE: “Neural overlap of L1 and L2 semantic representations in speech: a decoding approach”

AUTHORS: Eowyn Van de Putte<sup>1</sup>, Wouter De Baene<sup>1,2</sup>, Marcel Brass<sup>1</sup>, Wouter Duyck<sup>1</sup>

AFFILIATION: <sup>1</sup> *Department of Experimental Psychology, Ghent University, Ghent, Belgium*

<sup>2</sup> *Department of Cognitive Neuropsychology, Tilburg University, Tilburg, the Netherlands*

RUNNING HEAD: Neural representation of bilingual semantics

KEY WORDS: mvpa; production; neural overlap; fMRI; bilingualism

#### CORRESPONDING AUTHOR

Eowyn Van de Putte

Department of Experimental Psychology, Ghent University

Henri Dunantlaan 2, Ghent, B-9000, Belgium

Email: Eowyn.vandeputte@ugent.be

Tel: +32472281456

## **ABSTRACT**

Although research has now converged towards a consensus that both languages of a bilingual are represented in at least partly shared systems for language comprehension, it remains unclear whether both languages are represented in the same neural populations for production. We investigated the neural overlap between L1 and L2 semantic representations of translation equivalents using a production task in which the participants had to name pictures in L1 and L2.

Using a decoding approach, we tested whether brain activity during the production of individual nouns in one language allowed predicting the production of the same concepts in the other language. Because both languages only share the underlying semantic representation (sensory and lexical overlap was maximally avoided), this would offer very strong evidence for neural overlap in semantic representations of bilinguals. Based on the brain activation for the individual concepts in one language in the bilateral occipito-temporal cortex and the inferior and the middle temporal gyrus, we could accurately predict the equivalent individual concepts in the other language. This indicates that these regions share semantic representations across L1 and L2 word production.

Given the high prevalence of multilingualism in the world, the understanding of bilingual language processing is of high relevance for society. In the literature, bilinguals are defined as people who need and use two (or more) languages in their everyday lives (Grosjean, 1992), without necessarily being equally proficient in both languages. The last decennia, the study of bilingual language processing has rapidly gained interest in cognitive psychology.

Although there has been some debate to what extent the bilingual lexicon is integrated across languages, the three most influential behavioral models of bilingual language organization all assume that the semantic systems completely or partly overlap across languages. The Revised hierarchical model (Kroll & Stewart, 1994), with its focus on lexico-semantic links, and the BIA+ model (Dijkstra & van Heuven, 2002), with its focus on orthographic lexical representations, assume a shared semantic system. However, this does not imply that the meaning of every word should be completely identical in every language. Indeed, the distributed feature model (Van Hell & De Groot, 1998) assumes partially overlapping semantic features (instead of whole concepts) across languages, depending on specific characteristics of the concepts. Only the distributed feature model has focused in somewhat more detail on the organization of semantic representations and the factors that may influence it, such as concept/word concreteness. More specifically, Van Hell and De Groot (1998) argued that conceptual representations in bilingual memory depend on word-type and grammatical class. They found that the overlap in meaning, indexed by the number of shared features, is larger for concrete translations, cognates and noun translations, relative to abstract translations, noncognates and verb translations.

As in the behavioral literature, three main theories can also be discerned in the neuroimaging literature of bilingual language processing (Green, 2003; Paradis, 2004, 2009; Ullman, 2001, 2005). Although the behavioral models mainly focused on lexico-semantic representations, the neurally-based accounts consider syntax as well. Across the neural models, there is consensus about the lexico-semantic organization across languages, which is the focus of the present paper, but they mainly diverge with respect to syntactical representations. Ullman (2001, 2005) and Paradis (2004, 2009) both argue that with increasing proficiency the neural representation of second language syntax converges with the neural representation of L1 language syntax, whereas Green (2003) argues that already from the beginning of L2 learning, L2 syntactical representations recruit the same neural circuits as the L1 syntactical

representations. Overall, despite the substantive difference between these neural models, all three models make very similar assumptions and predictions and point in the direction of common semantic representations across L1 and L2 in high proficient bilinguals with an early age of L2 acquisition.

Despite the relative consensus among the neural models of bilingual language processing concerning lexico-semantic organization, the neuroimaging studies that investigated the hypothesis that the semantic systems of both languages are represented by overlapping, rather than distinct cortical language areas have provided very divergent results, probably due to their huge methodological heterogeneity. In these classical neuroimaging studies, the neural overlap between L1 and L2 semantic representations has been investigated using contrast designs in which an experimental condition is compared with a control condition. For instance, Illes et al. (1999) reported that semantic decisions activated different brain regions than non-semantic decisions, and then compared results between L1 and L2 words. Within these designs however, many studies used tasks in the experimental (semantic) condition that differed on phonological or orthographic processing demands and task difficulty, in addition to the targeted semantic processing demands (Binder, Desai, Graves, & Conant, 2009). For example, a semantic task like animacy judgment (e.g. *horse*: ‘is it living or nonliving’?) also relies on additional phonological and orthographic processes when comparing it with a control task that for instance involves nonword stimuli (e.g. nbgsj, nbqsj: ‘are they identical’?). Then, the comparison between L1 and L2 across such tasks may reveal the targeted cross-lingual semantic overlap, but also the overlap in the peripheral untargeted processing that may result from phonology, orthography, or even mere task difficulty, because the semantic tasks are often also more difficult than the control tasks that they are compared with (Binder et al., 2009). As such, the question about neural overlap of semantic representations across languages also needs to be assessed using other approaches.

Additional to this classical univariate approach, the fMRI-adaptation paradigm has been proposed as a useful tool to study the neural convergence between L1 and L2 representations in bilinguals (Chee, 2009). Adaptation refers to the phenomenon where the successive presentation of two identical stimuli elicits a smaller neural response than the successive presentation of two dissimilar stimuli. Neural overlap between the L1 and L2 semantic systems (e.g. Crinion et al., 2006) has been demonstrated with this approach. However, adaptation results are difficult to interpret

given its largely unknown neurophysiological underpinnings and its susceptibility to experimental demands, attentional confounds and novelty or mismatch effects especially for exact stimulus repetitions (e.g. Davis & Poldrack, 2013). Contrary to these univariate approaches, multi-voxel pattern analysis (MVPA) is sensitive to distributed neural representations and indexes a fundamentally different aspect of the neural code (Jimura & Poldrack, 2012; Epstein & Morgan, 2012). Compared to univariate activation or adaptation results, MVPA has been suggested to provide more direct measures of representations (Davis & Poldrack, 2013). MVPA cannot only detect that equivalent concepts have been presented in the two languages, but also that the representations of these specific concepts are similar across the two languages. MVPA distinguishes patterns of neural activity associated with different stimuli or cognitive states.

The logic of using this approach for the present purposes is that one uses a task in which the L1 vs. L2 response tap into very different orthographic, phonological, and sensory representations. Then the classifier may only predict the concept in one language based on the brain responses for the (translation) equivalent concept word in the other language if these two concepts in the different languages elicit similar semantic neural representations. If this is the case, this serves as direct evidence for the neural overlap of semantic representations in L1 and L2, supporting an integrative view of L1 and L2 in bilinguals.

In the literature, there are currently only 2 studies that used MVPA to investigate neural overlap of semantic representations in bilingual language processing (Buchweitz, Shinkareva, Mason, Mitchell, & Just, 2012; Correia et al., 2014), and both are situated in the language comprehension domain. Buchweitz et al. (2012) investigated the semantic representations tapped into by word reading (visual comprehension). Eleven proficient Portuguese-English bilinguals were asked to silently read concrete nouns from two semantic categories (tools and dwellings). Using MVPA, they could predict the individual nouns that the participants were seeing based on the neural representation of the equivalent nouns in the other language situated in the left inferior frontal gyrus, the left posterior superior temporal lobe, the postcentral gyrus and the left inferior parietal sulcus.

In the second study, Correia et al. (2014) focused on semantic representations in listening (auditory comprehension). Ten proficient Dutch-English bilinguals listened to concrete animal nouns and non-animal nouns in both languages and pressed a

button whenever they heard a non-animal word. They could accurately predict which animal noun was heard in one language based on the brain response of the equivalent noun in the other language. The shared representation across languages was situated in the left anterior temporal lobe, the left angular gyrus and the posterior bank of the left postcentral gyrus, the right posterior superior temporal gyrus, the right medial anterior temporal lobe, the right anterior insula and bilateral occipital cortex. Both studies provide evidence for the existence of common overlapping semantic representations across languages in comprehension, both in the visual and auditory domains.

Besides these two language comprehension studies, to our knowledge, no studies have used MVPA (or decoding) to investigate the neural overlap across languages of semantic representations used for language production (speaking) in bilinguals. In the behavioral literature, language comprehension and production are studied in mostly independent lines of literature, and some theoretical accounts assume different lexicons for production and recognition, and even between auditory and visual domains (Caramazza, 1997; Gollan et al., 2011; Roelofs, 2003). These separate systems are sometimes assumed to eventually contact a semantic system that is shared between modalities (Shelton & Caramazza, 2001). Some fMRI decoding studies supported this assumption: In a monolingual study, Simanova, Hagoort, Oostenveld, and Van Gerven (2014) investigated the possibility to decode the semantic category across modalities within L1. Participants had to perform a semantic categorization comprehension task with 4 types of stimuli (spoken words, written words, photographs and natural sounds) and subsequently produced the same stimuli afterwards in a free recall session. Simanova et al. (2014) found evidence for the shared representation of semantic information across input modality situated in the left inferior temporal cortex and frontal regions. Similarly, Van Doren, Dupont, De Grauwe, Peeters and Vandenberghe (2010) also reported overlapping neural semantic representations between the recognition of L1 words and L1 picture naming in the occipito-temporal regions and inferior frontal regions in a forced choice recognition task.

However, there's also evidence that semantic processing across comprehension and production might not rely on two completely overlapping semantic representations. Two other monolingual comprehension studies that investigated semantic processing showed different patterns of activation elicited by the passive viewing of pictures than

by the silent reading of the names of these pictures (Gates & Yoon, 2005; Reinholz & Pollmann, 2005). A possible explanation for this dissociation could be that names of pictures do not automatically activate the corresponding object-selective areas as pictures do.

In the present study, we will use a similar MVPA approach as Buchweitz et al. (2012) and Correia et al. (2014) used for respectively bilingual reading and listening (all comprehension), and Simanova et al. (2014) for monolingual language processing across modalities (production vs. comprehension). However, instead of looking at bilingual comprehension, we will examine bilingual production using a bilingual picture naming task. As such, this is also the first MVPA study to assess the neural overlap between the semantic representations that L1 and L2 production rely on.

## **MATERIALS AND METHODS**

### ***Participants***

Twenty-four right-handed individuals (12 males, 12 females; mean age = 23,38, range = 19-27 years) participated in the study. Fifteen participants were early French-Dutch bilinguals who acquired both languages from birth. Nine participants were late sequential bilinguals who learned French at school at the age of 9, as all children do in the Flemish educational system. The early bilinguals spoke French with their parents, Dutch at school and switched frequently between both languages with their friends. Three late sequential bilinguals followed an additional high level French language education program, two had a job in which they often had to speak both in Dutch and French and four only learned French at primary school, but rarely used it at the time of scanning.

The participants filled out a language background questionnaire to assess their subjective language proficiency, switching frequency and the age of acquisition of both languages. Additionally, proficiency in Dutch and French was measured with the LexTALE and the Boston Naming test (BNT; Kaplan, Goodglass, & Weintraub, 1983). The Dutch LexTALE (Lemhöfer & Broersma, 2012) that consists of 60 items and the French LexTALE (Brysbaert, 2013) that consists of 56 items are tests of vocabulary knowledge that give a good indication of general Dutch and French proficiency. The BNT is a 60-item picture naming test that measures word retrieval (see Table 1 for results on these proficiency measures).

All recruited participants had normal or corrected-to-normal vision. None of them used medication or had a history of drug abuse, head trauma, or neurological or psychiatric illness. All participants gave written informed consent before participating. The study was approved by the Ethical Committee of Ghent University hospital.

### ***Stimuli***

Pictures of 10 concepts had to be named in French and in Dutch. All stimuli were stored as  $720 \times 450$ -pixel images ( $18.1 \times 11.3$  visual degrees). Importantly, two completely different images were selected per concept (e.g. horse). Per participant, each image was associated with one language (for an example, see Figure 1). This image-to-language assignment was counterbalanced across participants. Visual similarities (e.g. point of view, colour) between the two images of the same concept and lexical overlap (overlapping phonemes and graphemes) between translation equivalents of the same concept were minimal. In order to avoid visual similarity, for each pair, both a black-white line drawing, and a color picture were used. Also, perspectives of the object were varied, such that no low-level visual features were shared across both pictures. The lexical overlap between translation equivalents of the same concept were quantified with the Levenshtein distance, in which the amount of insertions, deletions or substitutions required to change one word into the other is used as a measure of phonetic and orthographic distance (Levenshtein, 1965). The Levenshtein distance between the translation equivalents in Dutch and French was 1.00 for all stimuli, corresponding with a maximum number of changes, which equalizes a maximum orthographic and phonological distance between the Dutch and French translation equivalents. The translation equivalents were matched on word length ( $p=0.193$ ) and word frequency ( $p=0.885$ ). See Appendix for an overview of all experimental stimuli.

### ***Experimental design***

The neural overlap between Dutch and French semantic representations was examined using a production task in which the participants were asked to name the pictures in Dutch and French. This picture-naming task was organized in 2



consecutive parts (a Dutch and a French part). The order of the two language parts was counterbalanced across participants. Each language part included 7 blocks that always started with a familiarization phase to ensure picture-name agreement. To this end, each of the 10 pictures was presented on the centre of the screen with its name below it in the language relevant for the respective part. Participants had to press a button to proceed to the next stimulus. After this familiarization block, they worked through a practice block of 10 trials in which they had to name the 10 pictures, followed by 5 experimental scan blocks of 60 picture naming trials. These 60 trials included 6 randomised picture presentations of the 10 concepts. During each trial, one of the pictures was shown for 1000 ms, followed by a fixation screen of 1000 ms and a jittered stimulus onset asynchrony (mean = 2600 ms, range = 1000-5200 ms, in steps of 300 ms, distribution with pseudologarithmic density). At the start of each stimulus presentation, the naming was recorded during 3000 ms.

### ***Functional MRI data acquisition***

Participants were scanned with a 3T Magnetom Trio MRI scanner system (Siemens Medical Systems, Erlangen, Germany), using a standard 32-channel radio-frequency head coil. They were positioned head-first and supine in the magnetic bore and were instructed not to move their heads to avoid motion artefacts. The scanning procedure started for each participant with a high-resolution 3D structural scan, using a T1-weighted 3D MPRAGE sequence (TR = 2250 ms, TE = 4.18 ms, TI = 900 ms, acquisition matrix = 256 x 256 x 176, FOV = 256 mm, flip angle = 9°, voxels resized to 1 x 1 x 1 mm). After the structural images, whole brain functional images were collected using a T2\*-weighted EPI sequence, sensitive to BOLD contrast (TR = 2000 ms, TE = 28 ms, image matrix = 64 x 64, FOV = 224 mm, flip angle = 80°, slice thickness = 3 mm, distance factor = 17%, voxels resized to 3 x 3 x 3 mm, 34 axial slices). A fixed number of images (152) were acquired per run.

### ***fMRI data pre-processing***

SPM8 software (Wellcome Department of Cognitive Neurology, London, UK) was used for the preprocessing and data-analyses of the acquired fMRI-data. The first nine

scans of all runs were excluded from the analysis to minimize T1 relaxation artefacts. For each run motion parameters were estimated and runs with more than 15% of bad volumes were repaired by interpolation through the ArtRepair Toolbox v4 (<http://cibsr.stanford.edu/tools/ArtRepair/ArtRepair.htm>). Six runs in four different participants exceeded 15 % of bad volumes. A threshold of 1,5 % from the mean was used as criterion to categorize a volume as bad. From the 6 runs that were categorized as bad, 4 runs occurred in the L1 blocks and 2 runs occurred in the L2 blocks. The repaired motion regressors were used for all further analyses. The images were slice-time corrected and spatially realigned to their mean image by rigid body transformation. Additionally, the high-resolution structural image was co-registered with this mean image and normalized to the Montreal Neurological Institute (MNI) template. These normalization parameters were then applied to the functional images to ensure an anatomically-informed normalization. The time series data at each voxel were processed using a high-pass filter with a cut-off of 128 s to remove low-frequency artifacts.

The normalized but unsmoothed images were used to perform the multivariate decoding analyses to prevent the possible reduced sensitivity to extract the full information in the spatial patterns after smoothing. Therefore, smoothing was applied after the multivariate pattern classification analyses and prior to the second-level analysis using an 8 mm full-width half-maximum (FWHM) Gaussian kernel. Separately for the two language parts, statistical analyses were performed on individual subjects' data using the general linear model (GLM) in SPM8. All events were time-locked to the onset of the visual presentation. The fMRI time series data were modelled by 10 different vectors reflecting the semantic concept of the trial. All these vectors were convolved with a hemodynamic response function (HRF), as well as with the temporal derivative and entered into the regression model (the design matrix), which contained additional regressors to account for variance related to head motion. The statistical parameter estimates were computed separately for all columns in the design matrix.

### ***FMRI Data analysis: MVPA***

We performed multivariate decoding analyses with the PyMVPA toolbox (Hanke et al., 2009) to investigate the neural overlap between Dutch and French semantic representations in a production task. We employed a searchlight method (Kriegeskorte, Goebel, & Bandettini, 2006) to reveal local activity patterns that carry information about the semantic concept using a spherical searchlight with a radius of 3 voxels. Normalized but unsmoothed beta images were subjected to the analysis and a K Nearest Neighbours pattern classifier was used for classification. The use of other classifiers (The Gaussian Naïve Bayes classifier, the linear Support Vector Machines Classifier and the Radial Basis Function Support Vector Machines Classifier) yielded similar results. In each analysis, we used a leave-one-run-out cross-validation procedure. That is, for the across-language decoding analyses, the classifier was trained to discriminate between the activation patterns associated with the naming of each of the 10 concepts in one language for four of the five blocks (training data set). Subsequently, this pattern classifier was used to classify the activation patterns associated with the naming of the 10 concepts in the other language in the corresponding fifth block (test data set). Five-fold cross validation was achieved by repeating this procedure independently, with each block acting as a test data set once while the other blocks were used as training data sets. Classification accuracies were averaged across all five iterations, yielding a mean decoding accuracy map for each participant. These analyses were done in two directions: with Dutch trials as training trials and French trials as test trials and vice versa. The classifier was only able to accurately predict which concept was named if semantic representations of Dutch and French overlap in the brain. To assure that classifier performance only reflected the semantic overlap between the two languages, visual similarities between the two images of a concept and lexical similarities between the translation equivalents were maximally reduced. Additionally, we also ran within-language decoding analyses in which the training and test data were from the same language part. This by definition implied sensory overlap between pictures, contrary to the across-language analyses, which were our main focus and implied the use of different images of the same concept in the different languages to particularly exclude the visual confound in that specific comparison.

Classification accuracy significantly above chance (i.e.  $> 0.10$ ) implied that the classifier was able to accurately predict which concept was named, whereas chance level performance implied that it was not possible to predict the concept that was named. Note, however, that searchlight approaches can lead to interpretation errors such as the misidentification of a cluster as informative. For example, a cluster that is not informative can appear in the searchlight map if other clusters within the sphere provide significant classification accuracies (Etzel, Zacks, & Braver, 2013).

Therefore, to show that the significant clusters from the searchlight analyses are informative itself, cluster confirmatory analyses were additionally applied (Etzel, Zacks, & Braver, 2013). The main idea here is that the cluster should always be tested for information as a ROI, before describing it in any sense other than that of the centers of searchlights. If the ROI made from the cluster is informative, then there is justification for concluding that the cluster is itself informative (Etzel et al., 2013).

Additionally, evidence that the cluster contains the most informative voxels is provided if the global anatomically-defined area (defined on the basis of the AAL atlas) to which the cluster belongs but with the cluster voxels removed contains less information than the global area including the cluster and the cluster itself. If the area is still informative after the cluster has been deleted, the information should be described in terms of the area as a whole.

### ***Group analyses***

Whole brain, voxel-by-voxel second-level statistical analyses were performed to see how well decoding could be performed on average across all subjects (Haynes et al., 2007). The across-language decoding accuracies were averaged across the two directions (Dutch as training language and French as test language and vice versa). These resulting decoding accuracy maps were contrasted with chance level of accuracy (10%) using a one-sample t-test to reveal significant coding of semantic concepts across languages. Group maps significance was defined using a threshold of  $p < .001$  at voxel level and a cluster level corrected for the whole brain at  $p < .05$ .

The separate within-language decoding accuracy maps (same language (Dutch or French) as training and test language) were submitted to a flexible factorial design with language (Dutch or French) as within-subject factor. A disjunction analysis was

used to identify brain areas showing significant decoding accuracies in Dutch ( $p < .001$ ) but not in French ( $p > .05$ ) and vice versa. This analysis was done to investigate the brain regions that can discriminate between semantic concepts within Dutch, but not in French and vice versa. Note that these within-language disjunction analyses need to be interpreted with care, as within-language comparisons imply lexical overlap besides the semantic overlap. This makes it impossible to distinguish whether differences in the areas involved in the decoding within L1 or the decoding within L2 are due to differences in semantic representations or rather lexical representations.

Additionally, we performed region of interest (ROI) analyses on predefined ROIs. Based on the (monolingual) study of Simanova et al. (2014), we selected a number of candidate regions that we expected to be involved in semantic processing. In that study, a similar decoding approach was used to investigate the semantic processing in L1 during the presentation of pictures, written words, spoken words and sounds. We selected the brain regions that Simanova et al. (2014) reported to be involved in the semantic processing of pictures in L1, to see whether these regions also generalize to L2 (bilateral middle temporal gyrus, left fusiform gyrus, left middle occipital gyrus, right postcentral gyrus and right calcarine). Because of the similar approach that was used to investigate monolingual neural semantic representations, the study of Simanova et al. (2014) was very relevant as the base for the selection of the ROI's in our study to investigate the bilingual neural semantic representations.

Spherical ROIs (radius = 10 mm) were centered at the peak coordinates identified for each of these brain regions. To identify significant ROI regions the Bonferroni correction was applied.

### **Representational similarity analysis**

To test whether the classification can really be explained by semantic similarity, rather than visual similarity we additionally applied representational similarity analysis (RSA). To this end, we analysed the response similarities across languages between the evoked fMRI responses across all 10 stimulus pairs in the selected regions of interest (ROIs), based on the regions that we found in our whole brain analysis. To obtain the 10 x 10 similarity matrix for every ROI and for each subject,

we correlated the first level L1 beta images for all 10 stimuli with the first level L2 beta images for the 10 stimuli.

The RSA matrices for each ROI (similarity matrices between the brain responses evoked by the 10 stimuli in L1 and the brain responses evoked by the 10 stimuli in L2) were averaged across all subjects and correlated with a semantic similarity matrix of all 10 stimuli combinations and a visual similarity matrix of all the picture combinations using Spearman rank correlations (Kriegeskorte, Mur, & Bandettini, 2008). If the similarities of the brain activations across the 10 stimulus pairs correlated more with the semantic similarity matrix than with the visual similarity matrix, this provides additional evidence that the regions found in our whole brain analyses indeed reflect shared semantic and not higher-order visual processing, even though highly dissimilar pictures were used.

As a conservative approach towards our semantic processing claim, the semantic similarity matrix was drawn from an independent study, Snaut, a program that measures semantic distances between words (Mandera, Keuleers, & Brysbaert, in press). We used 1-semantic distance as a measure of semantic similarity. The visual similarity matrix was created based on subjective ratings of the visual similarity between all the combinations of pictures that were used in the experiment. The subjects that participated in this fMRI study had to respond on a 7-point Likert scale (1 = the pictures do not have any visual similarity, 7 = the pictures are visual identical).

## RESULTS

### *Whole brain statistical analyses*

#### Across-language decoding

To reveal significant coding of semantic concepts across languages, a one-sample t-test was used in which the decoding accuracy maps were contrasted with chance level (10%). For this analysis, the across-language decoding accuracies were averaged across the two directions (Dutch to French and French to Dutch). Significant across-language decoding accuracies were found in the left middle occipital gyrus extending into the left fusiform gyrus, the right lingual gyrus extending into the right inferior temporal gyrus and left inferior temporal gyrus extending into the left hippocampus (Figure 2; Table 2).

To show that the significant clusters from the searchlight analyses are informative itself, cluster confirmatory analyses was applied. For every significant whole brain searchlight cluster three ROI's were made: One ROI was created from the cluster itself, a second ROI was made from the global anatomically-defined area to which the cluster belongs and a third ROI was made from the global anatomically-defined area to which the cluster belongs but with the cluster voxels removed.

After cluster confirmatory analyses, all the clusters from the whole brain analyses were significant ( $p < .001$  for the cluster in the left middle occipital gyrus, the cluster in the right lingual gyrus:  $p < .001$  and the cluster in the left inferior temporal gyrus). Both the whole left middle occipital gyrus with the cluster ( $p < .001$ ) and the left middle occipital gyrus without the cluster were significant ( $p < .05$ ). However, the cluster alone contained more information than the brain area with the cluster ( $p < .001$ ) and the brain area without the cluster ( $p < .05$ ). This provides evidence that the information is widespread throughout the left middle occipital gyrus, with the most information found in the cluster centered at -39 -85 4. The whole right lingual gyrus with the cluster was significant ( $p < .05$ ) and the area without the cluster was not significant ( $p = .088$ ). The cluster alone contained more information than the brain area with the cluster ( $p < .001$ ) and the brain area without the cluster ( $p < .001$ ). This provides evidence that the cluster itself (9 -88 -2) contains the most informative voxels in the right lingual gyrus. The whole left inferior temporal gyrus with the cluster ( $p = .53$ ) and without the cluster ( $p = .58$ ) were not significant. The cluster

alone contained more information than the brain area with the cluster ( $p < .001$ ) and the brain area without the cluster ( $p < .001$ ). This provides evidence that the cluster (-42 -43 -26) itself contains the most informative voxels in the left inferior temporal gyrus. To conclude, all the clusters contained the most informative voxels, but the involvement in the left middle occipital gyrus was additionally more widespread.

### Within-language decoding

We also performed within-language decoding analyses to get a more fine-grained look at the regions that might be involved in the semantic processing of one specific language.

Disjunction analyses showed that the bilateral postcentral gyrus extending into the bilateral precentral gyrus, the left superior temporal gyrus, the right supramarginal gyrus, the right cuneus extending into the right superior parietal gyrus and the right middle temporal gyrus extending into the right inferior temporal gyrus were involved in L2 production, but not in L1 production (Figure 3, Table 3).

In the opposite direction, no significant decoding accuracies were observed for L1, that were not observed for L2. Note that, as mentioned above, this within-language disjunction analysis reveals cross-language differences, but do not allow to fully disentangle semantic from lexical involvement, given that within-language comparisons by definition also contain lexical (and visual) overlap.

Although this wasn't the primary goal of the study, for exploratory purposes, we also included AOA and proficiency as covariates to look at the differences between low and high proficient bilinguals and early and late bilinguals in the brain regions that showed significant decoding accuracies. Only the covariate AOA yielded significant differences between early and late bilinguals. An early age of acquisition of L2 correlates with lower decoding accuracies of L2 in the right calcarine, extending into the right middle occipital gyrus, the right cuneus and the left postcentral gyrus extending into the left inferior temporal gyrus, the left lingual gyrus, the left fusiform gyrus and the left inferior occipital gyrus (Table 4).



### ***ROI results***

In the ROI analyses, we selected the brain regions that Simanova et al. (2014) reported to be involved in the semantic processing of pictures in a first language to see whether these regions also generalize to a second language. After Bonferroni correction, the ROI's in the left middle temporal gyrus, the right middle temporal gyrus, the left fusiform gyrus, the left middle occipital gyrus and the right calcarine showed significant across-language decoding accuracies. Only the ROI in the right postcentral gyrus was not significant (Table 5).

### **Representational similarity analysis**

The RSA matrices of the three ROI's (Right lingual gyrus, Left inferior temporal gyrus, left middle occipital gyrus) correlated more with the semantic similarity matrix (ROI 1:  $r = .15$ ; ROI 2:  $r = .05$ ; ROI 3:  $r = .07$ ) than with the visual similarity matrix (ROI 1:  $r = .05$ ; ROI 2:  $r = .02$ ; ROI 3:  $r = .06$ ). For ROI 1, this correlation was significantly different, and for ROI 2 and 3 this correlation was not significantly different (ROI 1:  $p < .01$ ; ROI 2:  $p > .23$ ; ROI 3:  $p > .72$ ; paired-sample t-test).

## DISCUSSION

In the present study, we investigated the neural overlap between the semantic representations needed for L1 and L2 production, using multivariate decoding analyses. The results showed that significant decoding of individual concepts is possible across languages. Because lexical or sensory overlap was excluded across L1 and L2, the classifier could have only accurately predicted which concept was named in one language given the activation pattern for naming in the other language if semantic representations of L1 and L2 do overlap in the brain. These findings provide evidence for the existence of shared semantic representations that are situated in the bilateral occipito-temporal cortex and the inferior and the middle temporal gyrus. These regions align with monolingual studies that also situated (L1) semantic representations in the posterior temporal regions (Rodd, Vitello, Woollams, & Adank, 2015; Van Doren et al., 2010). Furthermore, these results indicate that when learning a L2, new lexical forms are mapped onto the existing areas that represent semantics for the existing (L1) language.

A point of discrepancy with previous (comprehension) studies (Binder et al., 2009; Buchweitz et al., 2012; Correia et al., 2014) is that for our production modality, we didn't replicate the involvement of frontal regions and anterior temporal regions in semantic processing. The dorsomedial prefrontal cortex, inferior frontal gyrus and the ventromedial prefrontal cortex showed up in a meta-analysis of 120 functional imaging studies that investigated the neural representation of the semantic system of spoken and written words in L1 comprehension (Binder et al., 2009). Across languages, the left inferior frontal gyrus showed shared semantic representations in visual comprehension (read concrete nouns in silence; Buchweitz et al., 2012) whereas the left anterior temporal lobe showed overlapping semantic representations across languages in auditory comprehension (listen to concrete nouns; Correia et al., 2014). This might indicate that the involvement of frontal regions and anterior temporal regions in semantics is more specific for comprehension than for production. To investigate whether neural overlap across languages is shared across modalities, future studies should investigate across-language semantic overlap in the different modalities within the same individuals. Another possible explanation for the absence of frontal structures in our paper should also be considered. The low selection demands and the overlearning of the pictures (through repetition) may explain the

absence of frontal structures in this task. Thompson-Schill, D'Esposito, Aguirre and Farah (1997) for example argue that frontal activation is involved in the selection of information among competing alternatives from semantic memory, but is therefore not the result of semantic retrieval per se. They argued that the involvement of the inferior frontal gyrus was absent or reduced in semantic tasks with low selection demands or high repetition. As such, the current picture naming task allows a more focused assessment of semantic processing, irrespective of irrelevant task demands.

In addition to the overlapping semantic representations across languages in the bilateral occipito-temporal cortex and the inferior and the middle temporal gyrus, we also found brain areas that showed significant decoding accuracies in L2, but not in L1. These results suggest that in addition to the shared neural populations representing semantics across languages, there are also neural populations that are recruited specifically by L2 at the semantic or lexical level (the bilateral postcentral gyrus extending into the bilateral precentral gyrus, the left superior temporal gyrus, the right supramarginal gyrus, the right cuneus extending into the right superior parietal gyrus and the right middle temporal gyrus extending into the right inferior temporal gyrus). The distinction between the semantic or lexical level is not possible to make in the disjunction of the within-languages decoding analysis, because only across-languages lexical overlap could be avoided in our design. The involvement of additional regions was more prominent in L2 than in L1, which suggest that the neural representation of a less proficient language is more widespread (Stowe & Sabourin, 2005).

Interestingly, the involvement of the neural populations in L2 semantic processing seems to be influenced by the AOA of L2. Our results seem to indicate that the later L2 was acquired, the more additional neural populations are involved in the semantic processing of L2. This might implicate a more efficient organization of conceptual knowledge in early bilinguals than in late bilinguals, as proposed in the reviews of Indefrey (2006), Perani and Abutalebi (2005) and Stowe and Sabourin (2005) who also suggested more extensive activations for L2 processing compared to L1 processing in late bilinguals, without dissociation between the specific modalities (e.g. comprehension and production). They concluded that late learners might be more likely to draw on additional resources to aid them in L2 processing. Note however, that we can't dissociate AOA and L2 exposure in this paper, because AOA is highly

correlated with the years of use of L2 in our sample. These effects could therefore be driven by both AOA or by the amount of exposure to L2.

Furthermore, we selected the brain regions that Simanova et al. (2014) reported to be involved in the decoding of the semantic category of pictures in L1 to see whether these regions also generalize to the semantic processing of pictures of individual concepts in L2. In the decoding across languages, the bilateral middle temporal gyrus, the left fusiform gyrus, the left middle occipital gyrus and the right calcarine were involved in our study. This finding again replicates the importance of the middle temporal gyrus not only for monolingual semantic representations (Price, 2012; Indefrey & Levelt, 2000), but also for common bilingual semantic representations in L1 and L2.

Despite the absence of low-level visual similarity between very dissimilar pictures of the same concepts, the representational similarity analysis for the left inferior temporal region and for the left middle occipital region seems to indicate that both visual and semantic features might have contributed to the classification. However, note that the RSA matrices of the three ROI's (Right lingual gyrus, Left inferior temporal gyrus, left middle occipital gyrus) correlated more with a semantic similarity matrix than with a visual similarity matrix<sup>1</sup>, even though semantic similarities were derived from an independent source (Mandera et al., in press). Secondly, Correia et al. (2014) also reported the involvement of occipital regions in a word listening task across languages, although no visual stimuli were used whatsoever. Therefore mental imagery could be a possible explanation in the sense that visual characteristics might be automatically activated during the (semantic) processing of concrete concepts (Binder & Desai, 2012). Thirdly, note that the other observed inferior and middle temporal regions are not typical reflections of visual involvement, but appear in previous monolingual meta-analyses as areas related to semantic processing (Price, 2012; Indefrey & Levelt, 2000).

Overall, the results of our study provide evidence for overlapping semantic representations of concrete concepts across L1 and L2 as suggested by all three theoretical models of bilingual language processing: the BIA + model, the revised hierarchical model and the distributed feature model (Dijkstra & van Heuven, 2002;

---

<sup>1</sup> Note that the difference between correlations was significant only for the right lingual gyrus, likely because of the (necessarily) small number of stimuli for which these correlations may be calculated.

Kroll & Stewart, 1994; Van Hell & De Groot, 1998). The distributed feature model, however, assumes less neural overlap for the semantic representations of abstract concepts across languages (Van Hell & De Groot, 1998). To test this assumption, future studies should compare the neural overlap in semantic representations of concrete and abstract concepts within the same individuals using a decoding approach. In the neuroimaging literature, our findings support Green's convergence hypothesis that also highlights the neural overlap between L1 and L2. More specifically, this theory assumes that during L2 acquisition, the neural representations of L2 will converge with the neural representations of L1 (Green, 2003). However, our findings also partially support Ullman's differential hypothesis (Ullman, 2001, 2005) and Paradis' neurolinguistic theory of bilingualism (Paradis, 2004, 2009). Although their focus is on the dissociation of neural areas that are used for L1 and L2, they also agree that with increasing proficiency, experience or an earlier age of acquisition, L2 representations might shift to rely more on the procedural structures of L1.

In our study we only saw an influence of AOA and not proficiency on the neural overlap. However future studies that specifically compare different (and therefore necessarily larger) subject groups with different AOA, proficiency levels and exposure levels are required to get a more detailed view on the influence of these individual difference variables on the neural overlap.

In addition to the influence of language use parameters (AOA, proficiency) it would also be interesting to look at the influence of language relatedness on the neural overlap of L1 and L2 semantic representations. Using an adaptation approach, Chee, Soon, & Lee (2003) for example investigated the neural overlap of semantic features across a more dissimilar language pair (Chinese – English) and reached a similar conclusion, namely that the Chinese and English semantic system have shared components, but also components that may be language-specific. Future MVPA research may systematically compare closer and linguistically/socioculturally more distant languages.

## ***Conclusions***

Brain activity in the bilateral occipito-temporal cortex and the inferior and the middle temporal gyrus associated with the activation of semantic representations of individual concepts during production in one language (e.g. “lune”) accurately predicts the activation of semantic representations of the equivalent concepts in the other language (e.g. “maan”). This suggests that these regions share semantic representations across L1 and L2 production. In addition, there are also brain areas that are recruited specifically by L2. These findings provide evidence for common, overlapping semantic representations.

## TABLES

**Table 1.** Overview of language proficiency scores (maximum score BNT: 60/ Lextale:100) for the simultaneous and sequential bilinguals. The self-ratings are on a 5-point likert scale and are summed across listening, speaking, reading and writing.

Group	Proficiency	Dutch (L1)	French (L2)
<b>Simultaneous bilinguals (<i>n</i>=15)</b>	Lextale	59.85 (6.96)	43.21 (21.30)
	Boston Naming Test	51.53 (5.22)	43.67 (6.04)
	Self-Ratings	19.53 (1.30)	17.93 (1.75)
<b>High level Sequential bilinguals (<i>n</i>=3)</b>	Lextale	64.99 (10.16)	61.31 (19.67)
	Boston Naming Test	56 (0)	41 (4.36)
	Self-Ratings	20 (0)	17.67 (2.52)
<b>Middle level Sequential bilinguals (<i>n</i>=2)</b>	Lextale	69.15 (1.20)	43.75 (16.42)
	Boston Naming Test	53 (1.41)	33 (2.83)
	Self-Ratings	20 (0)	15 (1.41)
<b>Low level Sequential bilinguals (<i>n</i>=4)</b>	Lextale	68.34 (3.04)	21.43 (3.57)
	Boston Naming Test	55 (2.94)	30.25 (7.85)
	Self-Ratings	20 (0)	13 (2.45)

**Table 2.** Results of the across-language decoding analyses. All thresholds were FWE corrected.

Brain region	X	Y	Z	z-score	Cluster size
Left middle occipital gyrus	-39	-85	4	5.25	635
Right lingual gyrus	9	-88	-2	4.94	773
Left inferior temporal gyrus	-42	-43	-26	3.85	113

**Table 3.** Disjunction: brain areas that showed significant decoding accuracies for L2 ( $p < .001$ ), but not for L1 ( $p > .05$ ). All thresholds were FWE corrected.

Brain region	X	Y	Z	z-score	Cluster size
Left postcentral gyrus	-54	-10	19	5.62	421
Right precentral gyrus	60	8	31	4.86	276
Right supramarginal gyrus	57	-22	37	4.93	142
Right cuneus	15	-67	40	4.79	269

**Table 4.** Mean accuracy L2 X covariate AOA L2. All thresholds were FWE corrected.

Brain region	X	Y	Z	z-score	Cluster size
Right calcarine	21	-79	10	4.47	379
Left postcentral gyrus	-27	-43	67	4.38	121

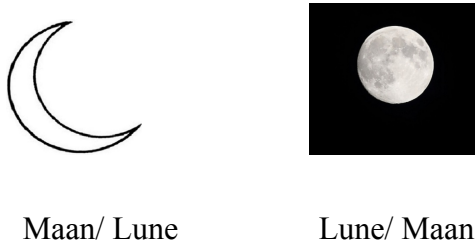
**Table 5.** Across languages Region of interest (ROI) analyses.

Brain region	X	Y	Z	Accuracy	P
Left middle temporal gyrus	-43	-63	10	0.1077	$p < .01$
Right middle temporal gyrus	48	-70	-1	0.1133	$p < .01$
Left fusiform gyrus	-40	-56	-15	0.1093	$p < .01$
Left middle occipital gyrus	-47	-81	-1	0.1170	$p < .001$
Right postcentral gyrus	45	-21	45	0.1040	$p = .13$
Left middle occipital gyrus	-22	-95	17	0.1089	$p < .01$
Right Calcarine	13	-98	3	0.1113	$p < .01$

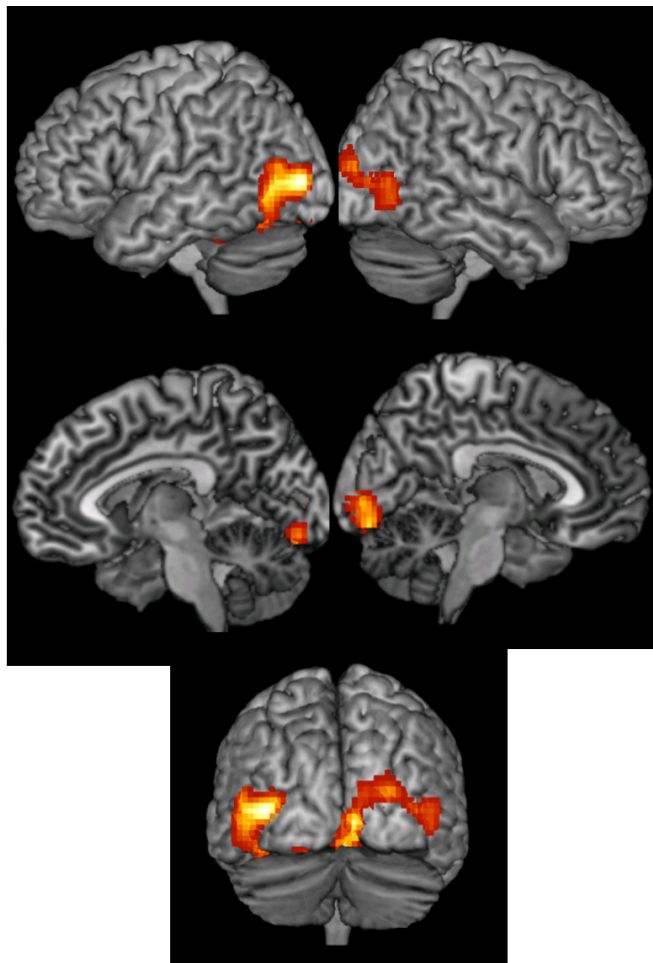


## FIGURES

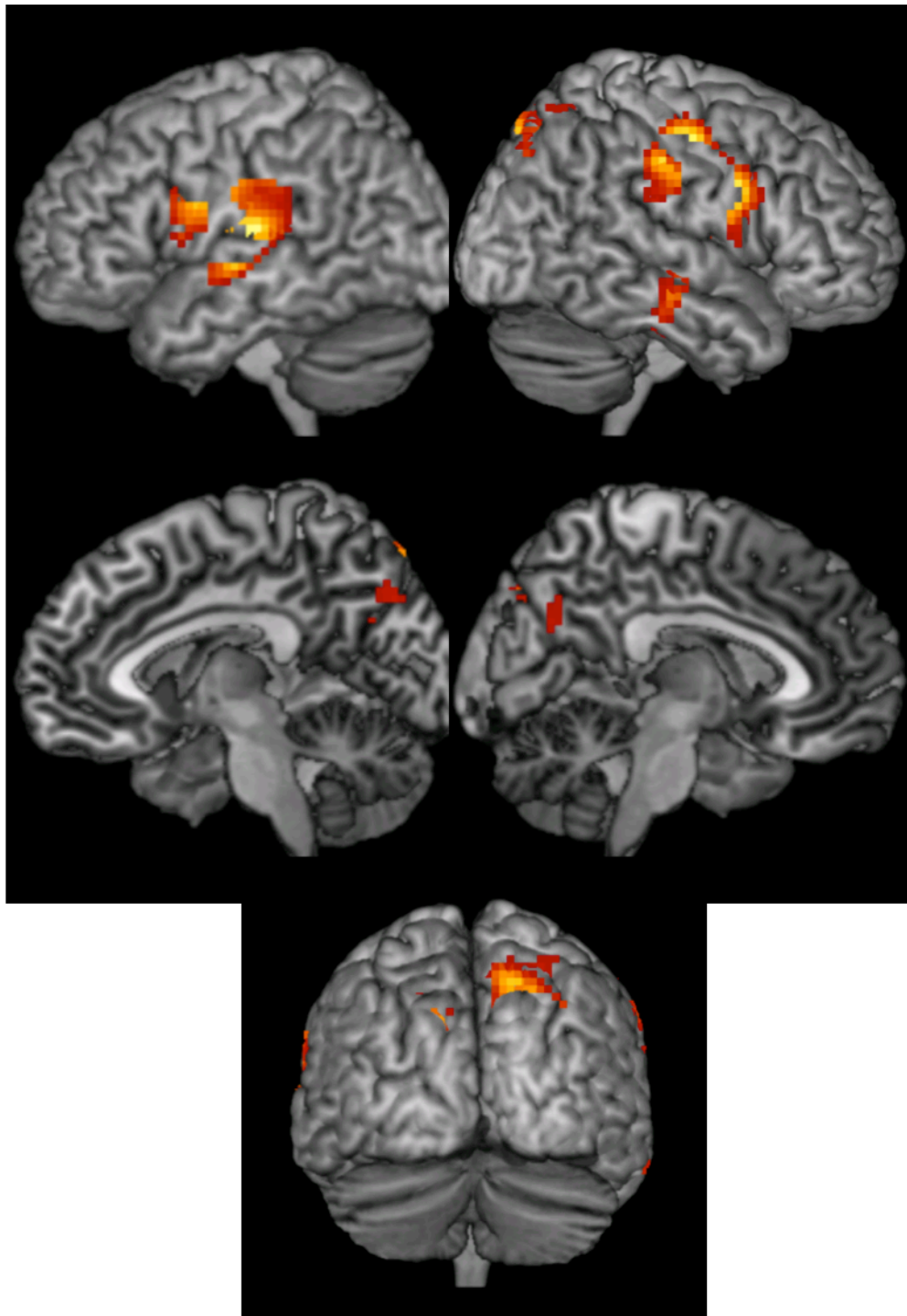
**Figure 1.** Pictures had to be named in French and in Dutch. For each concept (e.g. moon) two images with different visual features were selected, so that each language corresponded to a different picture.



**Figure 2.** Results of the whole brain searchlight analysis showing discriminability between semantic concepts in the generalization across languages. The color represents the t-values resulting from the group level analysis using a threshold of  $p < .001$  at voxel level and a cluster level corrected for the whole brain at  $p < .05$ .



**Figure 3.** Results of the disjunction analysis showing the brain areas that yielded significant decoding accuracies in L2 ( $p < .001$ ), but not in L1 ( $p > .05$ ).



## ACKNOWLEDGMENTS

This research was made possible by the Research Foundation-Flanders (FWO-Vlaanderen; FWO G058914N).









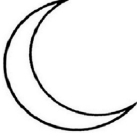

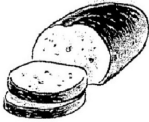







## REFERENCES

- Binder, J. R., & Desai, R. H. (2012). The Neurobiology of Semantic Memory, *15*(11), 527–536. <http://doi.org/10.1016/j.tics.2011.10.001>. The
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. <http://doi.org/10.1093/cercor/bhp055>
- Brysbaert, M. (2013). Lextale\_FR a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, *53*(1), 23–37. <http://doi.org/10.5334/pb-53-1-23>
- Buchweitz, A., Shinkareva, S. V., Mason, R. A., Mitchell, T. M., & Just, M. A. (2012). Identifying bilingual semantic neural representations across languages. *Brain and Language*, *120*(3), 282–289. <http://doi.org/10.1016/j.bandl.2011.09.003>
- Caramazza, A. (1997). How many levels of processing are there in lexical access?. *Cognitive Neuropsychology*, *14*(1), 177–208. <http://doi.org/10.1080/026432997381664>
- Chee, M. W. L., Soon, C. S., & Lee, H. L. (2003). Common and Segregated Neuronal Networks for Different Languages Revealed Using Functional Magnetic Resonance Adaptation. *Journal of Cognitive Neuroscience*, *15*(1), 85–97.
- Chee, M.W.L. (2009). fMR-adaptation and the bilingual brain. *Brain & Language*, *109*, 75-79
- Correia, J., Formisano, E., Valente, G., Hausfeld, L., Jansma, B., & Bonte, M. (2014). Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *The Journal of Neuroscience*, *34*(1), 332–338. <http://doi.org/10.1523/JNEUROSCI.1302-13.2014>
- Crinion, J. (2006). Language Control in the Bilingual Brain. *Science*, *312*(5779), 1537–1540. <http://doi.org/10.1126/science.1127761>
- Davis, T., & Poldrack, R. A. (2013). Measuring neural representations with fMRI: Practices and pitfalls. *Annals of the New York Academy of Sciences*, *1296*(1), 108–134. <http://doi.org/10.1111/nyas.12156>
- De Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1001–1018.
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. <http://doi.org/10.1017/S1366728902003012>
- Epstein, R. A., & Morgan, L. K. (2012). Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. *Neuropsychologia*, *50*(4), 530–543. <http://doi.org/10.1016/j.neuropsychologia.2011.09.042>
- Gates, L., & Yoon, M. G. (2005). Distinct and shared cortical regions of the human brain activated by pictorial depictions versus verbal descriptions: An fMRI study.

- NeuroImage*, 24(2), 473–486. <http://doi.org/10.1016/j.neuroimage.2004.08.020>
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis. *Journal of Experimental Psychology. General*, 140(2), 186–209. <http://doi.org/10.1037/a0022256>
- Green, D. W. (2003). The neural basis of the lexicon and the grammar in L2 acquisition. In R. van Hout, A. Hulk, F. Kuiken, and R. Towell (Eds.). *The interface between syntax and the lexicon in second language acquisition* (pp. 197–208). Amsterdam/Philadelphia: John Benjamins Publishing Company
- Grosjean, F. (1992). Another View of Bilingualism. *Advances in Psychology*, 83, 51–62. [http://doi.org/10.1016/S0166-4115\(08\)61487-9](http://doi.org/10.1016/S0166-4115(08)61487-9)
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1), 37–53. <http://doi.org/10.1007/s12021-008-9041-y>
- Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading Hidden Intentions in the Human Brain. *Current Biology*, 17(4), 323–328. <http://doi.org/10.1016/j.cub.2006.11.072>
- Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D. E., Glover, G. H., Poldrack, R., ... Wagner, A. D. (1999). Convergent Cortical Representation of Semantic Processing in Bilinguals. *Brain and Language*, 70(3), 347–363. <http://doi.org/10.1006/brln.1999.2186>
- Indefrey, P., & Levelt, W. J. (2000). The neural correlates of language production. In *The new cognitive neurosciences*; 2nd ed. (pp. 845–865). MIT press.
- Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning*, 56(1), 279–304. <http://doi.org/10.1111/j.1467-9922.2006.00365.x>
- Jimura, K., & Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, 50(4), 544–552. <http://doi.org/10.1016/j.neuropsychologia.2011.11.007>
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). Boston Naming Test (Revised). *Philadelphia: Lea & Febiger*.
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 3863–3868. <http://doi.org/10.1073/pnas.0600244103>
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. <http://doi.org/10.1006/jmla.1994.1008>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44(2), 325–343. <http://doi.org/10.3758/s13428-011-0146-0>
- Mandera, P., Keuleers, E., & Brysbaert, M. (in press). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*.

- Paradis, M. (2004). A neurolinguistic theory of bilingualism. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Paradis, M. (2009). Declarative and procedural determinants of second languages. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202–206. <http://doi.org/10.1016/j.conb.2005.03.007>
- Price, C. J. (2012). NeuroImage A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2), 816–847. <http://doi.org/10.1016/j.neuroimage.2012.04.062>
- Reinholz, J., & Pollmann, S. (2005). Differential activation of object-selective visual areas by passive viewing of pictures and words. *Cognitive Brain Research*, 24(3), 702–714. <http://doi.org/10.1016/j.cogbrainres.2005.04.009>
- Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: an Activation Likelihood Estimation meta-analysis. *Brain and Language*, 141, 89–102. <http://doi.org/10.1016/j.bandl.2014.11.012>
- Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken word forms. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* (pp. 115–159). Berlin, Germany: Mouton de Gruyter.
- Simanova, I., Hagoort, P., Oostenveld, R., & Van Gerven, M. a J. (2014). Modality-independent decoding of semantic information from the human brain. *Cerebral Cortex*, 24(2), 426–434. <http://doi.org/10.1093/cercor/bhs324>
- Stowe, L. a., & Sabourin, L. (2005). Imaging the processing of a second language: Effects of maturation and proficiency on the neural processes involved. *International Review of Applied Linguistics in Language Teaching*, 43(4), 329–353. <http://doi.org/10.1515/iral.2005.43.4.329>
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, 94(26), 14792–14797.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative / procedural model. *Bilingualism: Language and Cognition Language and Cognition*, 4(2), 105–122. <http://doi.org/10.1017/S1366728901000220>
- Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. *Mind and Context in Adult Second Language Acquisition*, 141–178.
- Van Doren, L., Dupont, P., De Grauwe, S., Peeters, R., & Vandenberghe, R. (2010). The amodal system for conscious word and picture identification in the absence of a semantic task. *NeuroImage*, 49(4), 3295–3307. <http://doi.org/10.1016/j.neuroimage.2009.12.005>
- Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1(3), 193–211. <http://doi.org/10.1017/S1366728998000352>

**Appendix.** Experimental stimuli. Overview of the 10 concepts that had to be named in Dutch and French and the two images that were selected per concept.

Dutch	French	Picture 1	Picture 2
Bed	Lit		
Bloem	Fleur		
Boom	Arbre		
Appel	Pomme		
Maan	Lune		
Brood	Pain		
Hond	Chien		
Glas	Verre		
Voet	Pied		
Paard	Cheval	