Bilingual word representation: a multi-task fMRI investigation

Laura Menenti[1]

Supervisor: Peter Indefrey[1,2]

[1]F.C.Donders centre for Cognitive Neuroimaging

[2]Max Planck Institute for Psycholinguistics

Address for Correspondence:

Laura Menenti

P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.

Tel.nr.: 024-3668493

E-mail: Laura.Menenti@fcdonders.ru.nl

Abstract

The Revised Hierarchical Model (Kroll and Stewart, 1994) assumes that words in a bilingual's languages have separate word form representations but shared conceptual representations. Two routes lead from an L2 word form to its conceptual representation: the word association route, where concepts are accessed through the corresponding L1 word form, and the concept mediation route, with direct access from L2 to concepts. The implications of the RHM for fMRI data were investigated. Proficient German-Dutch bilinguals carried out three tasks (voice decision, lexical decision and semantic decision) in their L1 and L2. LIFG and the left superior temporal gyrus were among the many areas responding more strongly to L2 than L1, while the right middle temporal gyrus was the only area responding more strongly to L1. Data from a previous behavioral experiment suggest that word association takes place in lexical but not semantic decision. An interaction between task and language was therefore expected in an area that processes L1 word forms. In a volume-of-interest analysis, such an area was found in the right posterior middle temporal gyrus. This area may therefore be involved in processing L1 word forms, in particular in the L1 processing that serves word association from L2.

Introduction

The representation of words in the different languages of bilingual speakers has been attracting a lot of interest from researchers in various disciplines over the years. In the field of cognitive psychology, the topic has been investigated for about twenty years (for a review, see French & Jacquet, 2004). In the field of cognitive neuroscience, the number of studies investigating neural representations of first and second languages has also been increasing (for reviews see Abutalebi, Cappa & Perani, 2001; Fabbro, 2001). Nevertheless, cross-talk between these disciplines has been scarce. It could be a fruitful approach however: so far fMRI researchers have found very varying results and have not been able to explain these in a coherent way, as we will see below. It is possible that cognitive models on bilingual language processing might shed light on when and why differences and commonalities between a bilingual's languages are found in brain data.

Neuroimaging evidence on the issue of word representation has been rather inconclusive. We will review a few of the studies here, considering only the direct between-language comparisons in these studies. These comparisons yield areas that are differentially active for the two languages, while not yielding any common areas. Therefore, such a comparison is only adequate to see to what extent two languages *differ*, not to what extent they share common areas.

First of all, a couple of studies investigating L2-L1 differences in language representation found no areas that consistently showed an effect of L1 vs. L2. The first study on bilingual language representation was done by Klein, Milner, Zatorre, Meyer & Evans (1995). In this PET study with English-French bilinguals, several word generation tasks in L1 and L2 were compared. These were word repetition, rhyme generation, synonym generation and translation. Of interest here, in a direct between-language comparison within tasks, no differentially active areas were

found. Illes, Francis, Desmond, Gabrieli, Glover, Poldrack et al. (1999) compared semantic (animal yes/no) and non-semantic (upper/lower case) decisions in first and second language. In this study the direct comparison between L1 and L2 did not reveal any consistent differences in the semantic condition (the direct comparison for the non-semantic condition was not reported). A problem of this study is that subjects were considered individually, not as a group. A second point of concern is that the study contained both English-Spanish and Spanish-English bilinguals. Language differences may therefore have clouded the L1-L2 difference. In an fMRI study with Spanish-English bilinguals, Pillai, Araque, Allison, Sethuraman, Loring, Thiruvaiyaru et al. (2003) also found no brain areas showing a language effect.

However, other studies have found differences between L1 and L2. Chee, Hon, Lee, & Soon (2001) also studied semantic and non-semantic decisions. The two tasks in this study were both matching-to-sample tasks with either English words or Chinese characters. In the semantic task, participants decided which of two probes was more similar in meaning to the target word, while in the non-semantic task they were instructed to choose the probe that was more similar in size on the screen to the target. The study had two groups of participants: English-Chinese and Chinese-English bilinguals. In the semantic task, both groups showed differences in the left anterior cingulate, the left middle frontal gyrus and the left frontal operculum in the L2-L1 comparison but no differentially activated areas in the L1-L2 comparison. For the L2-L1 comparison, the Chinese-English group also showed a differential activation in the left inferior temporal gyrus, while the English-Chinese group showed a difference in the right frontal operculum.

Ding, Perry, Peng, Ma, Li, Xu et al. (2003) also compared a semantic and a non-semantic task in L1 and L2. In a direct between language comparison, they found that the bilateral inferior parietal lobules and the bilateral post-central gyri responded more strongly to L2 than L1 in the

non-semantic task, while the right precentral gyrus and the right inferior frontal gyrus were differentially activated in the same comparison in the semantic task. The posterior left middle temporal gyrus responded more strongly to L1 than L2 in both tasks, while the left middle frontal gyrus only showed a language effect in the semantic task. In a PET-study with Dutch-French bilinguals De Bleser, Dupont, Postler, Bormans, Speelman, Mortelmans et al. (2003) found that L2 activated the left orbitofrontal gyrus and the cerebellar vermis more strongly than L1 while L1 activated the left middle frontal gyrus more strongly than L2. This latter finding is noticeable: while Ding et al. (2003) also found the left middle frontal gyrus to respond more strongly to L1 in their semantic task, Chee et al. (2001) found the left middle frontal gyrus to be more active for L2 than L1.

--------------------FIGURE 1 ABOUT HERE------------------

The evidence is therefore mixed: two studies found areas exhibiting stronger activation for L1 than L2, while four studies did not. Three studies found areas exhibiting stronger activation for L2 than L1, but three others did not. Moreover, there is not much agreement even among the studies that found the same difference: Ding et al. (2003) and De Bleser et al. (2003) find the left middle frontal gyrus (BA 44) to be more active for L1 than L2, while Chee et al. (2001) find this same area to be more active for L2. Ding et al. (2003) also find an L1-L2 difference in the left posterior temporal gyrus. In the L2-L1 comparison Ding et al (2003) find bilateral parietal areas to be activated, De Bleser et al. (2003) find the left orbitofrontal gyrus and the cerebellum, while Chee et al. (2001) find significant language effects in mainly posterior frontal areas. Due to these mixed results, it has so far been difficult to find a coherent interpretation for the activation patterns in the different studies. The present study is an attempt to use knowledge gained from cognitive

6

psychological research to interpret an fMRI experiment on bilingual processing. The behavioral model we base our research on is the Revised Hierarchical Model by Kroll & Schwartz (1994; see fig. 1). This model states that the words in a bilingual's two languages have shared representations at the conceptual level, but separate representations at the lexical level. There are both direct links between the L1 and L2 lexical stores and links from each of these stores to the conceptual level. However, these links are assumed to be asymmetric. The lexical link is assumed to be stronger from L2 to L1 than vice versa, while the conceptual link is stronger for L1 than for L2. To illustrate this, consider two types of translation: forward (L1-L2) and backward (L2-L1) translation. According to the RHM, forward translation mainly occurs by retrieving the concept belonging to the L1 word and then retrieving the L2 word form, while backward translation mainly occurs through direct association between the two word forms.

In a previous behavioral experiment (Menenti, 2006) we tested whether word association between L1 and L2 word forms indeed occurs. To achieve this we presented German-Dutch bilinguals with Dutch (L2) word pairs in which the translation equivalent of the first word rhymed with the second word (GRAP (joke) – *Witz* – FIETS (bicycle)). A phonological priming effect of the first word on the second would indicate that the L1 word form had indeed been activated. This mediated phonological priming effect was found in lexical decision but not semantic decision, indicating that word association takes place in lexical decision, while semantic decision is most likely achieved through the concept mediation route in this group of bilingual speakers.

From the previous experiment we conclude that L2 lexical decision will activate the L1 word form lexicon while semantic decision in L2 will not. The present fMRI experiment tries to build on this knowledge. In an auditory experiment, late proficient German-Dutch bilinguals performed three tasks in both their L1 and L2. These tasks were lexical and semantic decision and also voice decision, as a low-level task. Our hope was to disentangle common and differential activations

between languages in the following way: since the RHM predicts separate word form stores but shared conceptual stores between languages, activation differing between L1 and L2 should mainly be due to the word form level. Shared activations are more likely to be due to conceptual processing, as well as more general control processes due to the task at hand.

However, since L1 word forms are activated in L2 lexical decision, an interaction effect between task (lexical and semantic decision) and language is expected. An area that shows a stronger language difference in semantic than lexical decision might be related to L1 word form processing. In particular, such an area might be related to the part of L1 word form processing that is common to L1 and L2 in lexical decision but not semantic decision.

Method

Participants

14 German students from the University of Nijmegen and the Professional Institute of Arnhem and Nijmegen participated in this study. They were students of various disciplines, following their studies in Dutch. All had arrived in the Netherlands at least 10 months prior to the experiment and were following higher education in Dutch. Due to the great similarity between German (their native language) and Dutch and their intensive everyday use of Dutch, all participants were highly proficient in Dutch.

Stimuli

90 German-Dutch word pairs of translation equivalents were selected. The pairs all consisted of non-cognate items. The words were selected from the highest frequency words

8

(frequency > 40 occurrences / million) in the CELEX lexical databases for Dutch and German (Baayen, Piepenbrock & van Rijn, 1993).

For the lexical decision 90 nonwords were constructed by changing 1 – 3 letters in the 90 target words, with the restriction that the nonwords were phonologically legal in the target language. The stimuli were recorded in a soundproof booth by 4 speakers, 2 native Germans and 2 native Dutch (both 1 male and 1 female). As a low-level baseline, reversed speech stimuli were created by reversing the spoken words.

Design

The study consisted of three tasks: semantic decision (SD), in which participants decided whether a word corresponded to a living- or non-living concept, lexical decision (LD), in which participants decided whether an auditory stimulus was a word or a non-word and voice decision (VD), in which participants decided whether a word was spoken by a man or a woman. These tasks were all performed by pushing a button with either the index or the middle finger of the left hand. In all three tasks participants were instructed to press a button with the ring finger whenever they heard a reversed speech stimulus.

All tasks were performed in Dutch (L2) and German (L1). Task was blocked, with two different orders: SD-VD-LD and LD-VD-SD. Voice decision was always in the middle to minimize carryover effects from semantic decision to lexical decision and vice-versa. Language was varied in mini-blocks within the tasks. Blocks were 4-7 items long, and were always preceded by a visual cue (a Dutch or German flag) to inform the participant in which language the following items would be spoken. All tasks started with 3 warm-up items in Dutch and 3 in German, which were not considered in the analyses.

Procedure

The participants were given an instruction sheet in Dutch. After reading this they had the opportunity to ask questions. Once in the scanner, each task started with a cue reminding them of the task to be performed and the finger presses associated with the various responses. The experiment was performed in three separate runs, one for each task. The VD and SD runs lasted approximately 7 minutes, the LD lasted 14 minutes, due to the nonwords. The experiment ended with an anatomical scan, which lasted 9 minutes. A new trial started every 4.5 s. Stimuli were presented with a jitter of 1000s, the scanner was always silent while sounds were presented.

Scanning Parameters

Participants were scanned on a 3T Siemens Trio MR-scanner. To present the auditory stimuli, the scanning sequence contained a delay at the start of each volume. The TR was 4500 ms, of which the first 2500ms were always silent, allowing for jittered presentation of the auditory stimuli. Each volume contained 30 slices of 3.5mm thickness, with an inter-slice gap of 10% (0.35 mm). The TE was 30ms, the flip angle was 90º and the in-plane FOV was 224 x 224 mm^2.

Data Analysis

The data was analyzed using BrainVoyager QX software (Brain Innovation, Maastricht, the Netherlands). Intra-session alignment and motion correction were performed in one step by aligning all volumes of a subject to the first volume of the run before the anatomical scan. The functional data from all runs from all subjects were transformed to Talairach space (Talairach & Tournoux, 1988) and then combined in a fixed effects multi-subject GLM. Two runs had to be discarded, one because the subject had not performed the task correctly and one because the data was corrupt. Standardized reaction times were inserted as an extra predictor in all analyses, to

10

control for between-task differences in difficulty. A volume-of-interest (VOI) analysis was also performed on those areas that showed sensitivity to words compared to nonwords. In this VOI analysis beta values from all predictors from all VOIs were compared in a random effects analysis using SPSS.

Results

Whole-brain GLM

A whole-brain GLM was performed to see which areas in the brain showed sensitivity to the language of the stimuli. There was one predictor for each task x language condition (VD-D, VD-NL, LD-D, LD-NL, SD-D, SD-NL), two nonwords predictors (one for each language), one reversed speech predictor and an RT predictor. The visual cues preceding each block were also modeled with a separate predictor. The reversed speech stimuli were found to produce stronger activations than the word stimuli in most relevant brain areas. This was possibly due to the fact that these reversed words were intermixed within a task. When a participant encountered a reversed speech stimulus, he or she had to refrain from carrying out the task at hand and press a different button instead. Possibly, the participants even first unsuccessfully tried to continue performing their task. This may have caused strong responses to these stimuli. Since the reversed speech could no longer be regarded as a low-level baseline, this condition was not considered in the analyses.

-------------------TABLE 1 ABOUT HERE------------------------

Three contrasts were tested at the level of the whole brain: L2 words compared to L2 nonwords, L1 words compared to L1 nonwords and also a direct comparison between L2 words and L1 words. These contrasts were tested at False Discovery Rate (FDR) = 0.05, controlling for multiple comparisons. The results of this GLM are reported in table 1. Though there were areas that were activated both in the L1 words – nonwords and the L2 words – nonwords contrast (the bilateral posterior inferior frontal gyri, the left anterior middle and posterior superior frontal gyri, the bilateral posterior middle temporal gyri, the left anterior superior temporal gyrus, the right anterior and posterior cingulate gyrus and the left posterior cingulate gyrus), none of these were significant in a conjunction analysis on the L1 words - nonwords and L2 words - nonwords contrasts at FDR = 0.05. This means that even though these areas were involved in both L1 and L2 processing, they were more activated for one language than the other, but that this difference was not strong enough to be significant in the between language comparison.

In the direct comparison between L1 and L2 a number of areas were found that reacted significantly more strongly to L2 than L1. These were the right posterior superior frontal gyrus, the left anterior and posterior middle frontal gyrus, the left posterior inferior and medial frontal gyri and the precentral gyrus in the frontal lobe, the left mid superior and middle temporal gyri and the left posterior fusiform gyrus in the temporal lobe, the left posterior cingulate gyrus and the right insula. On the other hand, the posterior right middle temporal gyrus was the only area to respond more strongly to L1 than L2. Of the areas showing a language effect, the following areas were found to be activated in both the L1 words - nonwords and the L2 words - nonwords contrasts: the anterior middle frontal gyrus, the posterior inferior frontal gyrus, and the posterior cingulate, all in the left hemisphere, and the right posterior middle temporal gyrus. The right posterior superior frontal gyrus, left posterior middle and medial frontal gyri, left precentral gyrus, left mid superior and

middle temporal gyri the posterior fusiform gyrus and the right insula were therefore selective for L2. The right anterior middle temporal gyrus was selective for L1.

Task difficulty and RT

Task difficulty is a potential confound in our design. From our earlier behavioral experiment (Menenti, 2006) we know that RT's are slower in semantic decision than lexical decision. It is also likely that voice decision is an easier task than semantic or lexical decision. Moreover, it is probably easier to make lexical or semantic decisions about words in L1 than L2, because a speaker might not know all L2 words equally well. The RT data collected during the fMRI experiment is displayed in figure 2. As can be seen, both task and language seem to influence RT. A repeated measures 3x2 (task*language) GLM with RT as a dependent variable confirmed this. Both task and language had main effects on RT (task: $F_{i(2,88)}$ = 41.275, p =.000 ,$F_{s(2,10)}$ = 6.965, p = .013; language: $F_{i(1,89)}$ = 73.308,p = .000, $F_{s(1,11)}$ = 34.817 , p = .000). There was no interaction. This also indicates that the language of stimuli had an effect on RT's even in voice decision, even though in this task the meaning of the word was completely irrelevant. For the fMRI data it was necessary to control for difficulty by including reaction time (RT) as an additional predictor in all GLM's on the fMRI data. Variance due to RT is therefore discounted from the predictors of interest, task and language. Table 2 lists the areas sensitive to RT.

-------INSERT FIGURE 2 and TABLE 2 ABOUT HERE----

Of the areas mentioned above, five were found to also be sensitive to RT. These were the left posterior middle (which was significant in the L2-nonwords contrast) and inferior frontal gyri (significant in all three contrasts), the left mid superior temporal gyrus (significant in the L1 words-

13

nonwords contrast and the direct between-language comparison) and the right (significant in the between-language comparison) and left anterior insulae (significant in the L1 words-nonwords contrast). Although RT was a predictor in the GLM's described above as well, it is possible that not all variance due to RT was discounted from the effects found for language. It is therefore possible that some of the language effects mentioned above are still partly due to difficulty differences between L1 and L2.

Volume-of-Interest analysis

To take a closer look at the subtle effects we were looking for, a volume-of-interest (VOI) analysis was performed. VOIs were selected by using all the areas that showed a significant effect in either the L1-nonwords or L2-nonwords contrasts (see table 1). This was done because the purpose of this study is mainly to find areas serving the L1 and L2 lexicons. These areas should then be sensitive to whether a stimulus is a word or not. All areas with a significant result on one of the two tests were taken into the VOI-analysis. The advantage of this procedure was that areas were selected based on their sensitivity to words, not on the basis of sensitivity to language or task. Since these are effects of interest in our analyses, this would have unduly increased the probability of finding significant effects.

From the areas resulting from the words-nonwords contrasts, only those with more than 20 voxels were selected. This resulted in 29 clusters for the L2 comparison and 32 clusters for the L1 comparison. Of these, 8 coincided approximately. In these cases, the combination of the two was defined as one VOI. The total number of clusters in the VOI-analysis was therefore 53. The L2 words - nonwords and L1 words - nonwords contrast in table 1 list the clusters with their respective coordinates.

14

A random-effects 3x2 (task x language) repeated measures GLM was carried out on the VOI-data. Two clusters showed a significant task effect. These were the left posterior (talairach coordinates -6 -46 25[1]; $F_{2,10}$ = 5.155, p = 0.029) and right anterior (talairach coordinates 3 26 -2 $F_{2,10}$ = 4.725, p = 0.036) cingulate gyri. A main effect of language was found in four areas. These were the posterior right inferior temporal gyrus ($F_{1,11}$ = 5.758, p = 0.035), the posterior left middle temporal gyrus ($F_{1,11}$ = 5.741, p = 0.035), the left anterior insula ($F_{1,11}$ = 8.016, p = 0.016) and the right hippocampal gyrus ($F_{1,11}$ = 13.407, p = 0.004). Of these areas, the left insula had higher beta values for L2 than L1, while the other areas showed the reverse effect. A significant task*language interaction was found in the left postcentral gyrus ($F_{1,11}$ = 4.601, p = 0.038).

----------------INSERT FIGURE 3 ABOUT HERE ------------------

Our main hypothesis was that since L1 word forms are involved in L2 word processing in lexical decision but not in semantic decision, an area involved in L1 word form processing would be active in all L1 conditions, and also in L2 lexical decision but not L2 semantic decision. In other words, we predicted an interaction between task (lexical and semantic decision) and language, with an effect of language in semantic decision but not in lexical decision. In a 2x2 (task*language) repeated measures random-effects VOI-GLM three areas showed a significant task*language interaction. These were the right posterior middle temporal gyrus (talairach coordinates 39 -70 25; $F_{1,12}$ = 9.35, p = 0.01), the left postcentral gyrus (talairach coordinates -45 -25 49; $F_{1,12}$ = 10.91, p = 0.006) and the left posterior precuneus ($F_{1,12}$ = 6.61, p = 0.025). As can be seen, the left precuneus and the right middle temporal gyrus follow the pattern we predicted: there was no difference

---

[1] The talairach coordinates of all the clusters mentioned can be found in table 1. Where there is more than one cluster within the same region, talairach coordinates of the clusters are mentioned with the results.

between the two languages in lexical decision (L Pcu: $F_{1,12}$ = 0.004, p = .949; R GTm: $F_{1,12}$ = 0.329, p = .577) but a strong difference in semantic decision (L PCu: $F_{1,12}$ = 12.408, p = .004; R GTm: $F_{1,12}$ = 9.555, p = .009). For the left postcentral gyrus on the other hand, neither of the simple effects is significant.

-----------INSERT FIGURE 4 ABOUT HERE----------

Discussion

The basis of the revised hierarchical model (Kroll & Stewart, 1994) is the idea that there are two levels at which information on words is stored. The conceptual level is shared between languages, while at the word form level words are stored separately for a bilingual's languages. According to this model, brain areas that are common to L1 and L2 lexical processing are therefore more likely to be involved in conceptual processing, while areas that are different for L1 and L2 processing are probably areas related to word form processing. Of course, areas that are common to L1 and L2 processing can also reflect task-related processes, or more general processes such as listening. However, previous research can shed light on which of these tasks an area is more likely to accomplish. Moreover, in the present analyses common activations cannot be due to lower-order processes, because a relatively high-level baseline (phonologically legal nonwords) was used.

The GLM analysis revealed six areas that were involved in both L1 and L2 processing and did not show a significant difference between L1 and L2. These areas were the right posterior inferior frontal gyrus, the left posterior superior frontal gyrus, the left posterior middle temporal

16

gyrus, the left posterior precuneus and the right anterior and posterior cingulate gyrus. Frontal and cingulate areas are often found to be involved in attentional and strategic processes. It is therefore likely that these areas are common to L1 and L2 because they reflect task demands. The left posterior middle temporal gyrus has been found to be involved in passive listening to words, a condition in which there are no specific task demands (Indefrey & Cutler, 2004). The fact that this area appears to be common to L1 and L2 processing suggests that it is involved in semantic processing of words. However, none of the common areas were found to be significant in a conjunction analysis. Therefore, their activity was not equal for the two languages even though their activation differences did not reach significance in the direct language comparison.

There were two areas that were common to L1 and L2 but that did show a language effect. These were the right posterior middle temporal gyrus, which was more strongly activated for L1 than L2, and the left anterior middle frontal gyrus, which was more strongly activated for L2 than L1. The right posterior middle temporal gyrus has been found to be involved in passive word listening in 5 out of 10 studies that were considered in a meta-analysis by Indefrey & Cutler (2004). This area is therefore likely to be involved in storage of word-related information, rather than in task-related processing. The fact that it is more strongly activated for L1 than L2 suggests that it may be involved in processing L1 word forms. This area will be dealt with in more detail below. The left anterior middle frontal gyrus on the other hand is probably not involved in storage of word-related information but rather in strategic language processing. The fact that it is involved in both L1 and L2 processing but more strongly for L2 suggests that it subserves a component of lexical processing that is common to L1 and L2, but more effortful for L2. Since it shows differential activation to L1 and L2, this area is likely to be involved in word-form related processing in the context of judgment tasks.

There were then also areas that showed a language effect and were not common to the two languages. The biggest clusters were found in the left posterior inferior, middle and medial frontal gyri, the precentral gyrus, the left mid superior temporal gyrus and the right insula. All these areas responded more strongly to L2 than L1. This is partly in accordance with a recent meta-analysis by Indefrey (in preparation). In this study, the left posterior middle and inferior frontal gyri both showed stronger activation to L2 than L1 in at least two studies, as they did in our study. Note that contrary to the studies analysed by Indefrey (in preparation), the present study was done in the auditory modality. Activation differences between L1 and L2 in these areas therefore appear to be independent of presentation modality.

Of the areas showing an effect of stimulus language, the left posterior inferior and middle frontal gyri, the left mid superior temporal gyrus and the right insula were also found to be sensitive to RT. It is therefore likely that these areas were involved in processes related to task difficulty. The left posterior inferior frontal gyrus has also been found to be more strongly activated for L2 than L1 in a study by Xue, Dong, Jin & Chen (2004). In this study verbal working memory tasks were performed in L1 and L2, as well as control tasks. A greater activation for L2 than L1 in a difficult working-memory task again indicates that activations in this area might also be related to task difficulty. The left posterior middle frontal gyrus has been found to be more strongly activated for L2 than L1 processing before (Chee et al., 2001). This area has also been found to be active in L1 semantic processing of words and pictures. Our result then seems to be at odds with the idea that conceptual stores are shared between languages: if this is the case, there should be no difference between languages. As said, frontal areas are mainly thought to be involved in strategic language processing, not in storage of information. It is therefore not likely that the left posterior middle frontal gyrus contains the conceptual store, but rather that it reflects retrieval from this store. It is then in line with the predictions of the RHM that this retrieval is more effortful for L2 than L1: the

conceptual links are weaker for L2 than L1. The fact that the area is found to be sensitive to RT is in line with this idea.

Because the main aim of this study was to find lexical areas, a VOI analysis was carried out on those regions that showed sensitivity to lexicality of items. Areas that were significant either in the L1 words – nonwords or the L2 words – nonwords contrasts were selected for this VOI. Since these areas are sensitive to whether a stimulus is a word or not, they are likely to be involved in higher order word-related processes.

In this analysis, differences between L1 and L2 were found in four areas, the right posterior inferior temporal gyrus, the left posterior middle temporal gyrus, the left anterior insula and the right hippocampal gyrus. Of these, only the left insula showed a stronger response to L2 than L1, while the others had a stronger effect of L1. These four areas are both sensitive to language and to lexical status of items. Since they did not show an interaction with task, they are sensitive to language, irrespective of task. According to the RHM, they are therefore likely to be involved in word form processing. For the temporal areas, this is also in accordance with previous studies, claiming that lexical storage mainly takes place in the left temporal lobe (Indefrey & Cutler, 2004). Task effects were found in the right anterior and left posterior cingulate gyrus. These areas showed no effect of the language words were presented in. These findings are in accordance with the idea that the cingulate gyri are mainly involved in attention-related effects.

The RHM (see figure 1) claims that there are two routes from an L2 word to its corresponding concept: the word association route, in which an L2 word is directly associated to its corresponding word form and then to the concept, and the conceptual route, in which L2 words are directly linked to concepts, and to L1 words through concepts. In our previous experiment we found evidence for word association in lexical decision but not in semantic decision (Menenti, 2006). L1 word forms are therefore implicated in L2 processing in lexical decision but not in semantic

decision. This means that an area supporting this kind of L1 word form processing should show an interaction between task (lexical and semantic decision) and language, with an effect of language in semantic decision but not in lexical decision.

Such an interaction was found in three areas: the left precuneus, the right middle temporal gyrus and the left postcentral gyrus. Of these, the left precuneus and the right middle temporal gyrus showed the expected pattern, with a stronger difference between languages in SD and no difference in LD. According to our predictions, these areas may therefore be involved in L1 word form processing. For the right middle temporal gyrus this idea is confirmed by the whole-brain analyses, in which it was the only area showing a stronger activation for L1 than L2 words. Previous research also suggests that the right middle temporal gyrus is involved in L1 word comprehension. In a meta-analysis of 10 experiments comparing passive listening to L1 words with silence, the right posterior middle temporal gyrus was found to be activated in 5 of these studies (Indefrey & Cutler, 2004). On the other hand, in the present study this area was significantly more active for words than nonwords in both L1 and L2. This means that it is not exclusively involved in L1 processing. In sum, the right posterior middle temporal gyrus is involved in both L1 and L2 lexical processing, shows a stronger recruitment in L1 lexical processing and seems to support L1 access during L2 lexical decision.

Our claim is not that only the right middle temporal gyrus (and possibly the left precuneus) is involved in L1 word form processing. It is possible, in fact likely, that L1 and L2 word forms are largely processed in the same areas. Such areas would not be distinguishable from common areas involved in conceptual or task-related processing in our present analyses, because at the macroscopic level activation patterns overlap for L1 and L2. If this is the case, there are still two possibilities. It might be the case that at a more microscopic level L1 and L2 are distinct. Different neuronal population within the same area might be the memory store for L1 and L2 words. The

other option is that L1 and L2 words are to a large extent stored in the same neural networks. Cognitive models such as the BIA++ (Dijkstra & van Heuven, 2002) assume an integrated bilingual lexicon. Such models would therefore predict that L1 and L2 words are also stored within the same neural network. However, in the present study we do find at least one area that seems to be specifically sensitive to L1 word form processing. At least to some extent separate storage seems therefore likely. Moreover, there is also the option mentioned above: that even though L1 and L2 are partially represented in the same areas, they might still be located in different neuronal networks. A possibility to disentangle this issue is to carry out a more fine-grained analysis of the activation patterns in a particular brain area.

Another issue here is inter-subject variability. Even if language specific areas exist, they might be located in very different brain areas across different subjects. In this case a multi-subject analysis will not find many language-specific areas. A case in point here is a cortical stimulation study by Lucas II, McKhann II and Ojemann (2004). In this study the L1 and L2 sites of bilingual epileptic patients were mapped. Though almost all patients showed specific L1- and L2-dedicated sites, as well as shared sites, the distribution of these sites varied across patients. Many sites were therefore L1-specific in some patients but L2-specific in other patients. This study shows that L1- or L2-specific sites exist, but that they may not be the same across different brains.

The present study was an attempt to use knowledge gained from the field of cognitive psychology to make sense of the currently very mixed fMRI data. This is not an unproblematic enterprise. As some critics have said, cognitive models often do not make necessary predictions about the neural correlates of their components (Coltheart, in preparation). It is therefore difficult to test cognitive models with neural data. Nevertheless, the models do make predictions about which processes are distinct and which are not. It is at this level that fruitful collaboration between

cognitive and neural research can take place. As Beretta, Campbell, Carr, Huang, Schmitt, Christianson et al. (2003) claim:

"On the standard assumptions that the brain does the mind's work and that there is a one-to-one mapping between brain and mind, if our theory of mind distinguishes two processes, then the brain should distinguish two processes. Conversely, if our theory of mind says that there is no distinction between regulars and irregulars, then the brain should not distinguish them." (Beretta et al., 2003; p. 69.)

Though cognitive models do not say anything about which brain areas accomplish which tasks, they can therefore still be put to the test in the brain. In the present study we find at least one area that seems to be specific for L1 word form processing. Though it is not conclusive evidence, this fact fits better with models that postulate separate lexical stores, such as the RHM, than with models that assume an integrated lexicon, such as the BIA++. But apart from testing cognitive models using fMRI data, collaboration between cognitive psychology and neuroscience can also work the other way: interpreting brain data in the light of cognitive models will help in making sense of the very varying data.

22

References

Abutalebi, J., Cappa, S.F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition, 4(2),* 179-190.

Baayen, R.H., Piepenbrock, T. & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM).* Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Berretta, A., Campbell, C., Carr, T.H., Huang, J., Schmitt, L.M., Christianson, K., & Cao, Y. (2003). An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms. *Brain and Language, 85,* 67-92.

Chee, M.W.L., Hon, N., Ling Lee, H., Siong Soon, C. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage, 13,* 1155-1163.

Coltheart, M. (in preparation). What has functional neuroimaging told us about the mind (so far)?

De Bleser, R., Dupont, P., Postler, J., Bormans, G., Speelman, D., Mortelmans, et al. (2003). The organization of the bilingual lexicon: a PET study. *Journal of Neurolinguistics 16*, 439-456.

Ding, G., Perry, C., Peng, D., Ma, L.,Li, D., Xu, S. et al. (2003). Neural mechanisms underlying semantic and orthographic processing in Chinese-English bilinguals. *NeuroReport 14* (12), 1557-1562.

Dijkstra, A.F.J., & 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.

Fabbro, F. (2001). The bilingual brain: cerebral representation of languages. *Brain and Language, 79,* 211-222.

Fiez, J.A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human Brain Mapping, 5,* 79-83.

French, R.M., & Jacquet, M. (2004) Understanding bilingual memory: models and data. *TRENDS in Cognitive Sciences, 8(2),* 87-93.

Illes, J., Francis, W.S., Desmond, J.E., Gabrieli, J.D.E., Glover, G.H., Poldrack, R. et al. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language, 70* , 347-363.

Indefrey, P. (in preparation). Cerebral activation patterns in bilingual speakers and listeners.

Indefrey, P., & Cutler, E.A. (2004). Prelexical and lexical processing in listening. In M.S. Gazzaniga (Ed). *The Cognitive Neurosciences* (3rd ed.)*.* Cambridge, MA: MIT press.

Klein, D., Milner, B., Zatorre, R.J., Meyer, E.,& Evans, A.C. (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Science, USA 92*, 2899-2903.

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,* 149-174.

Lucas II, T.H., McKhann II, G.M., & Ojemann, G.A. (2004). Functional separation of languages in the bilingual brain: a comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients

Pillai, J.J., Araque, J.M., Allison, J.D., Sethuraman, S., Loring, D.W., Thiruvaiyaru, D. et al. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: preliminary findings. *NeuroImage 19*, 565-576.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Thieme.

24

Xue, G., Dong., Q., Jin, Z., & Chen, C.S. (2004). Mapping of verbal working memory in nonfluent

Chinese-English bilinguals with functional MRI. *NeuroImage, 22* (1), 1-10.

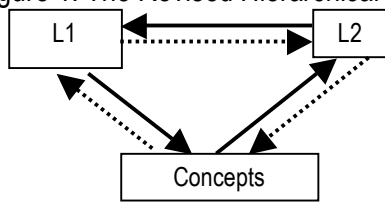Figure 1: The Revised Hierarchical Model (after Kroll & Stewart, 1994)

Table 1: clusters of activation in L1/L2 vs nonwords and L2-L1 contrasts

				L2 words - nonwords					L1 words- nonwords					L2 - L1				
				x	y	z	size	t	x	y	z	size	t	x	y	z	size	t
Frontal	R	posterior	GFs											36	32	31	19	3.588
			GFi	48	14	19	87	-4.08										
				39	20	10	31	-3.65	36	20	4	32	-3.64					
	L	anterior	GFs						-21	62	23	264	4.18					
			GFm	-30	47	25	340	4.27	-30	47	28	65	3.78	-36	47	22	10	3.873
			GFd						-3	50	7	843	4.73					
		posterior	GFs	-18	8	52	1194	4.83	-21	14	52	713	4.56					
			GFm	-36	-4	52	352	4.50						-51	17	31	259	4.469
				-42	17	37	92	4.11										
				-42	26	28	47	3.86										
			GFi											-45	11	4	3383	5.18
			44	-36	8	19	31	-4.16	-48	8	13	9	-3.72					
			45						-54	26	4	73	-4.07					
			GFd											-6	5	49	1223	4.988
		motor	ventralGPrc											-45	-1	34	689	4.53
			dorsalGPrc											-51	-10	47	278	4.171
														-27	-10	56	119	4.335
Temporal	R	anterior	GTm											53	0	-26	72	-3.64
		mid	GF						26	-37	-17	363	3.87					
			GH						30	-34	-5	95	3.79					
		posterior	GTs	54	-43	19	414	4.24										
				-45	-49	28	157	4.43										
			GTm	39	-58	10	567	4.50	30	-52	10	598	4.02	36	-58	19	129	-4.3
									39	-70	25	119	4.03	48	-61	19	77	-3.8
			GTi						45	-67	1	401	4.10					
	L	anterior	GTs	-45	20	-23	58	4.25	-64	-46	13	40	-3.83					
		mid	GTs						-57	-10	1	8	-3.82	-66	-22	7	529	5.579
			GTm	-64	-34	-14	101	4.11						-63	-7	-8	15	3.649
		posterior	GTm	-39	-70	25	432	4.37	-39	-70	16	447	4.95					
				-48	-64	16	485	4.93										
			GF											-45	-49	-17	85	4.403
Parietal	R	sensory	ventralGPoC						57	-22	16	676	4.50					
									51	-7	16	195	4.12					
		anterior	LPi	54	-40	37	176	4.84										
		posterior	FOF						30	-34	16	170	4.16					
	L	sensory	dorsalGPoc						-45	-25	49	66	4.08					
									-30	-28	49	44	4.02					
		anterior	LPi						-39	-25	28	138	4.14					
			LPs						-33	-40	49	56	3.70					
		posterior	LPs	-24	-52	53	84	4.23										
			PCu	-6	-52	46	911	4.01	-3	-52	49	116	4.00					
				-24	-52	52	29	4.22	-12	-49	37	1011	5.01					
Occipital	R	lateral	Go						24	-82	16	108	3.82					
Cingulum	R	anterior		3	26	-2	26	3.74	18	38	13	176	4.62					
				0	20	-8	33	4.05	6	20	-8	67	4.33					
		posterior		9	-34	34	100	3.97										
				9	-55	19	238	4.39	9	-52	13	95	4.07					
	L	anterior		-6	50	7	1387	4.79										
				-12	29	-5	35	3.86										
				-6	17	-8	47	4.23										
		posterior		-9	-34	31	182	4.00	-12	-58	19	600	4.85	-6	-61	16	38	3.988
				-6	-46	25	1566	5.13										
Insula	R	anterior												27	15	7	751	4.644
	L	anterior							-39	17	4	13	-3.68					
		posterior							-30	-7	16	82	4.38					
Cerebellum	R	medial							12	-52	-11	133	4.194					
		lateral		52	-46	-23	54	4.075										
	L	medial																
		lateral		-45	-49	-23	305	4.914										

Figure 2: RT per task and language
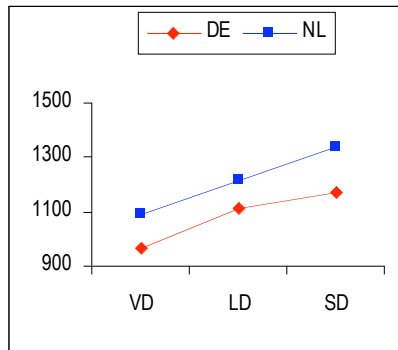
Table 2: clusters sensitive to RT (FDR = 0.05)

				RT				
				x	y	z	size	t
Frontal	R	posterior	GFm	42	11	31	8	3.31
			GFi	44	25	20	251	3.27
				48	25	7	409	3.93
	L	posterior	GFs	-6	23	49	396	4.74
			GFm	-39	8	31	1616	5.06
			GFi (BA44)	-51	16	16	601	4.15
			GFi(46)	-45	32	7	1156	4.82
Temporal	R	mid	GTs	45	-16	10	5566	5.74
				57	-28	10	233	3.94
		posterior	GTm	45	-43	4	152	4.01
	L	mid	GTs	-60	-22	1	8599	5.61
			GTi	-55	-61	-5	164	3.92
		posterior	GTm	-45	-43	7	3436	5.79
Parietal	R	posterior	Lpi	42	-49	43	349	4.28
	L	anterior	LPs	-30	-55	53	108	4.10
		posterior	PCu	-30	-70	40	297	4.11
Cingulum	L	anterior		-9	8	46	184	4.01
Insula	R	anterior		27	17	4	411	4.95
	L	anterior		-33	14	4	41	3.62

Figure 3: Task * Language interaction effects in left precuneus (A), right middle temporal gyrus (B) and left postcentral gyrus (C).
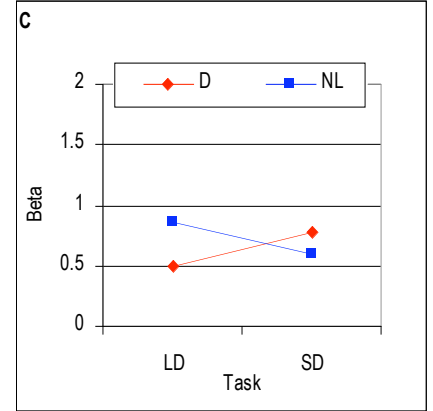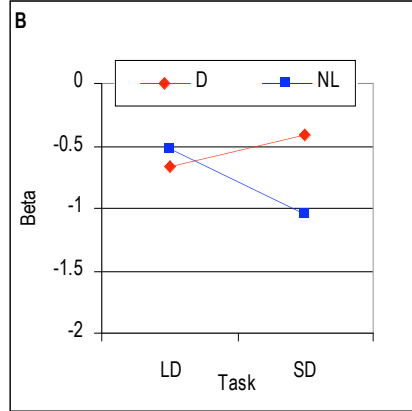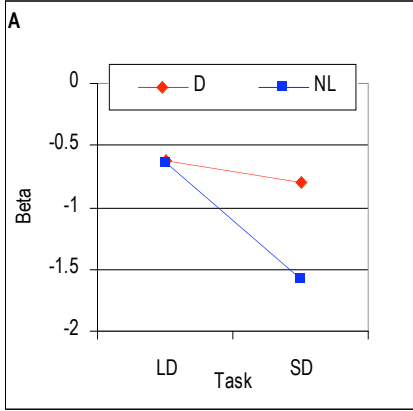
Figure 4: Region of interest in right posterior middle temporal gyrus