The Neural Representation of Concepts in Bilinguals: An Evaluation of Factors Influencing Cross-language Overlap Using fMRI-based Multivariate Pattern Analysis

eman ta zabal zazu

UPV EHU

Author:
Usman Ayub Sheikh

Supervisors:
David Soto
Manuel Carreiras

March 2, 2023

The Neural Representation of Concepts in Bilinguals: An Evaluation of Factors Influencing Cross-language Overlap Using fMRI-based Multivariate Pattern Analysis

eman ta zabal zazu

UPV EHU

DISSERTATION

Submitted in Partial Fulfillment of

the Requirements for

the Degree of

DOCTOR OF PHILOSOPHY (Cognitive Neuroscience)

at the

UNIVERSIDAD DEL PAÍS VASCO (UPV)
BASQUE CENTER ON COGNITION, BRAIN & LANGUAGE (BCBL)

by

Usman Ayub Sheikh

March 2, 2023

Acknowledgements

First of all, I thank Allah Almighty for blessing me with the strength and patience to complete this arduous work.

Words cannot express my gratitude to my supervisor Dr. David Soto for his invaluable patience and feedback. He has been an ideal mentor, offering advice and encouragement with a perfect blend of insight and humor. I also could not have undertaken this journey without continuous guidance and feedback of my co-supervisor, Dr. Manuel Carreiras.

Thanks to BCBL for allowing me just the right atmosphere and culture for academic as well as extracurricular growth and success. I am indebted to the best fMRI lab team ever at BCBL, for their magnanimous share of knowledge and expertise, especially Cesar Caballero and David Carcedo.

I am also grateful to my colleagues at BCBL, for thoughtful discussions and feedback during project presentations and lab meetings, especially Dr. Liv Hoversten, Dr. João Correia and Dr. Brendan Costello.

Thanks should also go to the administration team at BCBL, especially Eider Juaristi and Ana Fernandez, who helped with all the paperwork related to research grant and residency in Spain, and continued having patience with me until the end.

This endeavor would not have been possible without the generous support from the Spanish Ministry of Economy and Competitiveness (MINECO), who financed my research through the 'Severo Ochoa' Programme.

Lastly, I would be remiss in not mentioning my family, my parents, spouse, and children. Their belief in me has kept my spirits and motivation high during this process. I would also like to thank my three cats, Tomás, Sofi and Tigre for all the entertainment and emotional support.

Usman Ayub Sheikh

March 2, 2023

to *my beloved parents*

ABSTRACT

The Neural Representation of Concepts in Bilinguals: An Evaluation of Factors Influencing Cross-language Overlap Using fMRI-based Multivariate Pattern Analysis

by

Usman Ayub Sheikh

Supervisors:
Prof. David Soto, Ph.D.
Prof. Manuel Carreiras, Ph.D.

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Cognitive Neuroscience)

March 2, 2023

The neurocognitive mechanisms that support the generalization of semantic representations across different languages remain to be determined. Current psycholinguistic models propose that semantic representations are likely to overlap across languages, although there is evidence also to the contrary. Neuroimaging studies observed that brain activity patterns associated with the meaning of words may be similar across languages. However, the factors that mediate cross-language generalization of semantic representations are not known. In a series of functional MRI research studies, we investigate how factors including state of visual awareness, depth of word processing and lexico-semantic characteristics of words influence cross-language generalization of semantic representations. Using multivariate pattern analysis, we found that fully conscious and deep processing of high concrete and high frequency words leads to above-chance cross-language generalization in putative areas of the semantic network. These results have ramifications for existing psycholinguistic models and theories of meaning representation.

Nomenclature

MRI	Magnetic Resonance Imaging
fMRI	Functional MRI
MVPA	Multivariate Pattern Analysis
BOLD	Blood Oxygen Level Dependent
SVM	Support Vector Machine
L1	First Language of a Bilingual
L2	Second Language of a Bilingual
DICOM	Digital Imaging and Communication On Medicine
NIfTI	Neuroimaging Informatics Technology Initiative
MB	Multiband
TR	Repetition Time
FSL	fMRI Software Library
FEAT	fMRI Expert Analysis Tool
BET	Brain Extraction Tool
ICA	Independent Component Analysis
PCA	Principal Component Analysis
AROMA	Automatic Removal of Motion Artifacts

ROI	Region of Interest
IPL	Inferior Parietal Lobe
ATL	Anterior Temporal Lobe
LTL	Lateral Temporal Lobe
VTL	Ventromedial Temporal Lobe
FFG	Fusiform Gyrus
PHG	Parahippocampal Gyrus
dmPFC	dorso-medial prefrontal cortex
IFG	Inferior Frontal Gyrus
vmPFC	Ventromedial Prefrontal Cortex
PCG	Posterior Cingulate Gyrus
FLIRT	fMRI Linear Image Registration Tool
FWHM	Full Width at Half Maximum
ANOVA	Analysis of Variance
SD	Standard Deviation
FDR	False Discovery Rate
SOA	Stimulus Onset Asynchrony
BIA	Bilingual Integrated Activation
HFHC	High Frequency High Concreteness
LFLC	Low Frequency Low Concreteness

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Theories of Meaning Representation

How is the semantic information related to words and objects represented in the human brain? This is a key fundamental question that has motivated multidisciplinary research efforts into the most fundamental problems related to the structure of the mind and the brain. This line of research has culminated in the discovery of several important empirical phenomena and a vast number of theoretical positions.

Initially, neurophysiological studies of patients with lesions and more recently, non-invasive neuroimaging methods including electroencephalography (EEG), magnetoencephalography (MEG), transcranial magnetic stimulation (TMS) and magnetic resonance imaging (fMRI) have turned out to be instrumental in this effort. Below I provide an overview of the main questions addressed:

1. How is conceptual knowledge organized in the brain?
2. How different regions/sub-regions in the visual processing pathway contribute to conceptual processing of words and objects?
3. How different categories of words/objects might be separately represented in distinct regions/sub-regions?

Research groups looking into neural specificity for different categories of concepts made a significant number of observations pointing to the engagement of sensorimotor areas in conceptual processing. This drove interest in the idea that the conceptual content is distributed in the sensorimotor systems, also referred to as the *embodied cognition hypothesis* [4]. The most important difference between the embodied framework and previously existing theories was its emphasis on the format (modality-specific or independent format) rather than the nature of the conceptual content. The central principle is that the same neural systems that are involved in coding the perceptual properties and motor response of a concept are also involved in the representation and processing of that and related concepts in the semantic memory [4, 5, 6, 7, 8]. This stands in radical contrast to the view that the concepts are represented in a modality-independent system separate from the mechanisms involved in their perception and motor organization, often referred to as the *dis-embodied view*[1] [10, 11, 12, 13, 14]. On the other hand, the findings emerging from the patient data showed a slightly different picture. Specifically, the general result was that the lesions in the sensory motor regions have a minimal impact on the conceptual processing. This motivated arguments that brain lesions may not affect the processing of a given concept, if conceptualization processes can change the degree of its reliance from modality-specific information to other brain regions. These theories, often referred to as the *hybrid accounts* of cognition, suggested that while the processing of a given concept may involve the relevant sensorimotor systems, the center of gravity of this processing may not be confined to these systems, but distributed across many systems [15, 16].

This thesis contributes three different functional magnetic resonance imaging studies of Spanish-Basque bilinguals that attempt to address some of the unresolved issues using well-known multivariate pattern decoding methods, while making further indications regarding how to continue on with this line of inquiry. But before these experiments are presented and their findings discussed, I will start by setting up the theoretical context. In the following few subsections I explain further the three viewpoints i.e. embodied, dis-embodied and hybrid accounts in the light of existing evidence and literature.

[1] This term has been used by Guy Dove [9] and others.

1.1.1 Embodied Theories

Also referred to as the grounded cognition and situated cognition, these theories suggest that the semantic representations are grounded in the sensory-motor or modality-specific systems instead of being represented in an amodal system [4, 5, 6, 7, 8]. According to these theories, conceptual representations are not arbitrary and are determined by the content of the concept i.e. sensory concepts are represented in a sensory format and action concepts in a motor format. Among well-known embodied theories are: sensory/functional theory [17, 18, 19], convergence zones [20, 21], perceptual symbols systems [4, 8, 22] and neural reuse [8].

People suffering from neurodegenerative diseases like Alzheimer or Semantic Dementia have been observed to selectively lose knowledge of some semantic categories e.g. living animate entities or nonliving things. In an attempt to explain such category-specific deficits, Warrington and her colleagues proposed an idea referred to as the *sensory/functional theory* [17, 18, 19]. This theory suggests that the knowledge of a specific semantic category is located near the sensory-motor areas of the brain involved in the perception of its perceptual qualities and execution of its related movements. Therefore, when a sensory-motor area is damaged, the processing of instances of the specific semantic category that rely on that area is damaged. According to this theory, there is a high correlation between certain semantic categories and modality specific systems e.g. the identification of animals depends on the visual property of motion while the identification of fruits depends on the visual property of color. Neuroimaging research confirms this by showing different neural activation for different semantic categories e.g. Chao et al. 1999 [23] and 2002 [24] showed differential activation for animals and tools. Similarly, Kanwisher et al. 1997 [25] showed neural specificity for faces. Other studies have also shown differences in either the level of activation or the loci of activation for places, body parts, written words, tools and actions.

A related but different explanation was proposed by Damasio 1989 [20] and Damasio and Damasio, 1994 [21] which then came to be known as the *convergence zone theory*. According to this theory, visual perception of an entity entails activation of certain feature detectors (features being color, orientation and direction of movement etc.) in the relevant sensory-motor regions. The patterns of activation that arise in different sensory-motor areas are then stored in different association areas referred to as convergence zones (CZs). These CZs are hierarchically organized with those near visual processing areas storing activation patterns in the visual system and those near the motor areas storing activation patterns in the motor system. In the same way, higher-level CZs link together patterns from different modality-specific areas. Different CZs thus specialize in representing different semantic categories e.g. a CZ linking the shape and action features might be more important for the knowledge of tools while the CZ linking shape and movement might be more important for the knowledge of living things. Importantly, the main role of the CZs is to reactivate previously active patterns corresponding to a concept in relevant feature detectors, which then basically form the representation of that concept. CZs themselves do not hold a stand-alone representation of any concept (see also [22]).

Another important theory in the embodied framework is that of *perceptual symbol systems theory* [4, 8, 22]. According to this theory, a concept is a perceptual symbol that represents neural activation patterns that arise during its perception. These symbols are unconscious as they function unconsciously, they are componential as they can be built up from simpler parts arranged hierarchically, they are schematic as they are based on a schematic picture of a perceptual experience, and can therefore be more abstract than details of shape, position and proportion [26]. Attentional mechanisms are known to affect and shape perceptual symbols [22]. Since perceptual symbols are patterns of neural activity, their activation is flexibly adapted depending on the context and the goals of the perceptual experience. In the same way, while the symbols already stored allow recognition of related concepts, they can be modified, refined and updated through further experience. Additionally, the perceptual symbols created in different modalities can be grouped together depending on their nearness and co-occurrence in the semantic space (e.g. big and small) as well as the real world experience (e.g. table and chair). Once a group of perceptual symbols are linked and stored in the memory, they constitute a long-term memory.

Barsalou [8] presents another important idea i.e. *neural reuse*: the neural areas involved in the perception and action related to a concept are also involved in the processing of that concept. More specifically, through a perceptual simulation, the configurations of neurons established during interaction with a specific object are re-enacted. In this way, neural circuits established for the purpose of perception of a concept are recycled and put to use during conception, identification and processing of that concept. The theory further supports the mechanism by which the stored perceptual experience schemas can be combined productively to create more complex simulations [22].

The principal neuroimaging evidence for all these theories comes from the observation that conceptual processing usually leads to the activity in modality-specific regions of the brain [27, 28, 29, 30, 31]. However, while most of this evidence clearly demonstrates the activation of neural regions previously associated with sensory-motor information and shows the distributed nature of semantic representations, it falls short of proving that the conceptual representations themselves are made up of sensory-motor components [32]. It is also worth noting that fMRI studies exploring non-sensory-motor and abstract concepts have also localized their effects to brain regions overlapping with well-known sensory-motor areas. This comes off as a surprise as it is difficult to comprehend how sensorimotor representations can capture the content of abstract concepts. So, while these embodied accounts explain well how we represent concrete concepts, they cannot be directly extended to abstract concepts [33].

1.1.2 Dis-embodied Theories

Also known as amodal symbolic view [22], these theories suggest that concepts are represented in a modality-independent conceptual domain, disconnected and remote from the mechanisms of sensory and motor organization [12]. There exist systematic connections between amodal symbols and corresponding sensorimotor representations and depending on the task, the activation spreads from amodal symbols to sensorimotor representations.

The *symbolic model* of mind, also referred to as the *computational theory* of mind, describes the mind as a symbol system, characterized by a set of arbitrary symbols (atomic or composite) manipulated based on certain explicit rules [10, 11, 12, 13, 14]. The symbols are referred to as amodal because they are inherently non-perceptual. This theory derives its strength from its ability to explain certain phenomena such as productivity and systematicity of conceptual processing. Several findings from neuropsychological studies also support the amodal format of concepts. For example, in semantic dementia, brain damage in the temporal pole and the surrounding areas is known to impair conceptual processing [15]. And this degradation has been found to span across all conceptual categories and modalities.

Additionally, the abstract relations between concepts and semantic generalizations are also believed to require a single amodal semantic hub [34].

1.1.3 Hybrid Theories

Also known as the pluralistic view [16, 35], these theories suggest that some concepts (e.g. concrete nouns) are grounded in sensory-motor representations while others (e.g. abstract concepts) are amodal or dis-embodied [15]. Besides, the main role of semantic memory is to generalize across different concepts that have similar semantic significance and yet different sensory-motor attributes. If semantic memory was made up of just the modality-specific attributes of things then it is not clear how the higher-order generalizations can be achieved. These theories can also explain why lesions in some perceptual regions lead to uni-modal, category-specific deficits and in others as anterior temporal lobe lead to multi-modal, category-general semantic deficits [15, 36]. Among two well-known hybrid theories are the *distributed-plus-hub view* and the *conceptual topography theory*.

Distributed-plus-hub view proposes that sensory-motor information is necessary but not sufficient for conceptual processing and there is a need for a single hub that supports the interactive activation of representations in all modalities and for all semantic categories [15]. This is different from the previously mentioned convergence zone theory [20, 21] in a number of different ways with the most notable being that the convergence zone theory hypothesizes the existence of more than one specialized hubs, each encoding differently the association between different attributes/features (shape, corresponding actions) of a concept. In contrast, the single hub in the distribution-plus-hub view is supposed to encode all associations between different pairs of attributes/features of a concept. These representations are hypothesized to be amodal i.e. they can come from any perceptual/sensory modality and can be used to lead behavior in any expressive/motor modality. On the other hand, any damage to the hub is expected to lead to an impairment independent of the sensory-motor modalities. The principal evidence for this theory comes from neuropsychological studies. Lesion studies showed anterior temporal lobe (ATL) to be very important for semantic processing of famous faces and animals [37]. Functional as well as structural imaging studies have further shown those suffering from semantic dementia to have lesions in the anterior temporal lobe [38, 39, 40]. It is worth noting though that the significant ATL activation observed using PET is found to be largely absent in fMRI studies. One reason suggested for ATL's this shyness to fMRI is MRI's so-called susceptibility artifact [15].

Conceptual topography theory (CTT) [41], another well-known hybrid approach, aims to combine amodal accounts with perceptual theories inspired by recent studies of category-specific deficits. Specifically, it modifies Damasio's convergence zone theory [20]; so, while Damasio's convergence zones (CZs) themselves play no role in the representation of concepts, CTT proposes that CZs can also play representational roles, specifically during processes such as categorisation of familiar objects. So, during categorisation of known objects e.g. table/chair, active feature detectors feed activation into relevant CZs that can then integrate the table/chair-related features. These CZs can in turn feed activation to an expressive system e.g. one that could vocally transmit the word table/chair. It is worth noting that in this process of categorisation, the pattern of neural activation formed at CZ level is sufficient for correct transmission of category. In other words, reactivation of feature detectors is not required. However, if the process is more demanding, e.g. if it involves manipulating the corresponding conceptual representation, reactivation of a relevant feature detector pattern will become necessary. CTT is made up of four subsystems; so, each sensory-motor modality has feature detectors, analytic CZs, holistic CZs and modality CZs. Feature detectors for detecting and representing low-level features e.g. color, shape etc., analytic CZs for integrating modality-specific features, holistic CZs for integrating holistic conceptual properties (e.g. eyes, nose etc.). Modality CZs then integrate both analytic and holistic conceptual properties. In addition to these, there are cross-modal CZs that then integrate modality-specific CZs. Fernandino et al. 2015 [42], an fMRI-based MVPA study, investigated activation associated with 5 sensory-motor attributes (e.g. color, shape) for 900 words and found corresponding neural activation patterns to reflect multimodal abstraction. In general, the pattern of activation generated by different low-level and high-level attributes was found to be consistent with CTT.

1.2 Overlap of L1 and L2 Meaning Representations in Bilinguals

Visual word processing entails activation of a number of different processes, these range from orthographic operations at low-level visual areas (posterior areas including left occipitotemporal region and superior temporal gyrus) to semantic processes in mainly high-level association areas (a left-lateralized network of seven regions including inferior frontal gyrus, dorsomedial prefrontal cortex etc.) of the brain [1]. Given the fact that there are so many people who speak more than one language now, it has become very important to investigate how multiple languages are organized. In this thesis, we focus our research on bilinguals that use Spanish and Basque in their day to day life while having a satisfactory level of proficiency in both languages. Some of the participants of our studies had acquired both languages at the same time, others acquired one language first and the other later. A key unresolved question is whether different languages in bilinguals are integrated in the same system with shared/overlapping representations or rely on separate systems/representations for each language. Behavioral evidence from cross-language priming studies suggests that semantic representations are at least partially overlapping [43, 44, 45]. The evidence has led to the development of psycholinguistic models of bilingual language representation [46, 47]. Although these models differ in their predictions about the mechanisms that underlie lexical processing and the links between lexical and semantic processing of the two languages, they agree that semantic representations are at least partially overlapping between languages. Yet, other studies have failed to support overlapping semantic systems [48, 49, 50].

1.2.1 Behavioral Findings

Most of the primed lexical decision tasks have demonstrated facilitation of target words of one language preceded by semantically related primes of another language [43, 51]. These cross-language priming effects were found to be stronger from L1 to L2 than vice versa and for concrete than abstract pairs [45, 52]. Similarly, in semantic categorization tasks, where the participant had to decide whether the second word presented is a member of the category indicated by the first word, response times were found to be equivalent for word pairs from same/different languages [51, 53, 54]. In the word association task where the participant sees a word in one language and is supposed to say as fast as possible a semantically-related word in another language, Van Hell and DeGroot [47] found the retrieval of associate word to be easier for concrete as compared to abstract, for cognates as compared to non-cognates and for nouns as compared to verbs. Interestingly however, Francis and Goldmann [55] found similar cross-language priming effects for both abstract and concrete words, demonstrating a perfect overlap in semantic representations across languages independent of the level of concreteness.

On the other hand, some behavioral studies have failed to find evidence supporting overlapping semantic systems. De Groot and Nas [49] failed to show significant cross-language priming. Similarly, other studies showed the semantic representations to be more connected to one language than the other [48, 50].

1.2.2 Psycholinguistic Models

Based on the behavioral evidence related to semantic representations across languages, different psycholinguistic models explaining bilingual semantic representations and processing were proposed. Among well-known models are: the *revised hierarchical model* [46], the *distributed feature model* [47] and *bilingual interactive activation model (BIA+)* [56]. Since most of the behavioral findings support at least partially overlapping semantic representations across languages, all of these psycholinguistic models assume partially or completely overlapping semantic representations.

The revised hierarchical model was proposed by Kroll and Stewart in 1994 [46]. This model focuses on asymmetric lexico-semantic links. While assuming different lexical representations for each language, it assumes common semantic representations across languages. It was basically used to model the interaction between lexical and semantic representations during translation from the first language (L1) to the second language (L2) and vice versa. During acquisition of L2, word forms are often learned by associating them with corresponding translations in L1. Therefore, this model assumes that there are stronger links between L1's word forms and the language-independent semantic representations. In other words, L1 to L2 translation is expected to engage more the semantic representations than L2 to L1 translation. The model predicts however that as proficiency in L2 increases, the links between L2 word forms and the semantic representations strengthen which in turn improves the reliance of L2 to L1 translation as well on the semantic system.

The distributed feature model was proposed by Van Hell and de Groot in 1998 [47]. This model hypothesizes separate lexical stores but partially overlapping semantic representations across languages. The overlap in semantic representations is assumed to depend on the types of concepts, the individual and the cultural context in which the concepts get learned and processed. In this way, this is the only psycholinguistic model that explains in more detail the organization of semantic representations and the factors that may influence it including the level of concreteness of the concept being represented. Specifically, Van Hell and de Groot found the cross-language overlap in the semantic representations to be larger for concrete words, cognates and nouns (vs. abstract words, non-cognates and verbs).

The bilingual interactive activation model (BIA+) was proposed by Dijkstra and Van Hueven in 2002 [56]. While this model assumes common semantic representations across languages, it hypothesizes that the word forms are still stored in different lexicons for each language. It proposes an integrated non-selective access view; consequently, word candidates in both languages are activated in parallel and stored in an integrated lexicon. According to this model, a written word activates its lexical and sublexical orthographic and phonological representations which in turn activate the semantic representations and language nodes that indicate membership to a specific language.

1.2.3 Neuroimaging Evidence

Functional magnetic resonance imaging (fMRI), a non-invasive brain imaging method with a high spatial resolution but a low temporal resolution, is used to map neural activity associated with different brain functions based on corresponding blood oxygenation level dependent (BOLD) response. One analysis approach often used for the fMRI data involves treating all the voxels (each of an array of elements of volume that constitute the 3D fMRI image pspace) within a region of interest to be similar. An average of activation over all the voxels within that region, compared with the baseline activity, is thus considered to represent the brain activity or response corresponding to a specific task or condition. This approach is referred to as the classical univariate analysis [57]. Another related approach has proven to be effective in the study of convergence between neural representations corresponding to L1 and L2 of bilinguals. As the name suggests, this approach considers adaptation i.e. the successive presentation of two dissimilar stimuli leading to a smaller neural response as an indicator of neural overlap between the stimuli [58]. An important limitation of univariate activation-based approaches is that they are not best suited to identify whether or not semantic processing is mediated by a similar system across L1 and L2. Importantly, the observation that a cortical area is activated in both languages does not imply that the patterns of activity or neural representations are also similar across languages. This is an important motivation behind multivariate/multivoxel pattern analysis (MVPA) approaches. Whereas univariate approaches take the average of multivoxel activation to represent the overall activation within a region, MVPA considers the multivoxel activation patterns themselves to reflect the representational content. It is expected that if semantic representations overlap across languages, it would be possible to decode/classify the corresponding concepts across languages. MVPA has previously been shown to provide more direct measures of representations, to be sensitive to distributed neural representations and to classify patterns neural activity corresponding to different stimuli [59, 60, 61, 62, 63].

1.2.3.1 Univariate FMRI Studies

Majority of studies investigating the neural overlap across the two languages of the bilinguals have shown no difference in the semantic activation between the two languages [64, 65, 66, 67, 68]. This cross-language neural overlap was however found to be influenced by language proficiency. For example, for low-proficient bilinguals, Chee et al. 2001 [69] and Xue et al. 2004 [70] reported stronger activation in L2 as compared to L1 in the left posterior inferior parietal lobe, left anterior cingulate gyrus, left posterior middle frontal gyrus and the left posterior inferior frontal gyrus. This and other factors e.g. age of acquisition and exposure make it hard to generalize the findings from one study.

1.2.3.2 Multivariate Pattern Analysis Studies

Univariate approaches can be used to localise areas of the brain with common activation across languages. But this common activation can result from either different neural populations representing different languages or same populations representing different languages. It is only in the latter case that the integrative view of L1 and L2 can be proved. This is an important limitation of univariate approaches and a motivation behind adopting multivariate pattern analysis (MVPA) approach. MVPA approach, specifically, cross-language generalization, involves training the decoder to classify conceptual categories in one language and using it to predict the translation-equivalent concepts in the other language. In this way, if a decoder trained to classify between animals and tools in L1 within a specific area of the brain can identify animals and tools in L2 within same area of the brain with an above-chance classification accuracy, overlapping neural populations within that area of the brain can be proved. Instead of taking the mean of activation within a region to reflect the neural representation of a concept, MVPA looks at the multivariate pattern information within that region. This is why, compared to univariate and neural adaptation approaches, MVPA has been shown to provide more direct measures of neural representations [59, 60, 61, 62, 63].

Two recent studies used MVPA based cross-language decoding to assess whether the brain activity patterns elicited by words in one language can predict the patterns of equivalent words in the other language [71, 72]. They found language-shared representations in well-known semantic substrates including the left parietal lobe, inferior frontal gyrus, and posterior temporal lobe. An important limitation of these studies however is that the factors underlying the cross-language generalization of semantic representations across languages remain to be determined.

1.3 Outline of this Dissertation

Research on the representation of meaning in bilinguals has looked into language production and comprehension processes using different experimental paradigms based on both behavioral and neural activity measurements. However, most of these studies fell short in: controlling for confounding processes like automatic stimulus-response mappings (and other orthographic and phonological process) and differentiating between co-activation of neural regions vs. language-shared representations. On the one hand, the diversity of paradigms and bilingual populations ensured generalizability of results across these studies. On the other hand, it made it difficult to reach conclusions regarding neural overlap of semantic representation across languages. Recent fMRI studies (see section 1.2.3.2) succeeded in controlling for a number of these factors using MVPA, specifically, cross-language decoding analysis. The results obtained were promising and allowed for localization of language-shared semantic representations in previously well-known semantic areas of the brain. It remained to be determined however that which factors underlie this cross-language generalization of meaning. To provide novel insights into different potential task-related and lexico-semantic factors mediating generalization of semantic representation across languages while controlling for untargeted confounding processes, we proposed three different fMRI studies, investigating semantic processing in Spanish-Basque bilinguals, using cross-language semantic category decoding.

In the first empirical chapter (**chapter 2**), the factor of conscious visual awareness was considered. Specifically, an fMRI-based MVPA study of Spanish-Basque bilinguals, using animal/non-animal Spanish and Basque translational equivalents as stimuli, was used to investigate whether the semantic category (animal/non-animal) of partially conscious or non-conscious words can be decoded from multi-voxel patterns of activity in putative semantic areas of the brain. Secondly, using cross-language decoding, it was seen how different levels of awareness (fully-, partially- and non-conscious) affect cross-language generalization of semantic category in the considered semantic areas of the brain. Because the levels of awareness were controlled for using subjective ratings as well as objective performance measures, and the Spanish/Basque stimuli were controlled for linguistic properties (e.g. length, frequency and cognateness etc.) across languages and categories, this study offers strong evidence related to the influence of levels of awareness on the semantic representation in general and language-shared representations in particular.

In the second empirical chapter (**chapter 3**), the factor of depth of processing was considered. Specifically, an fMRI-based MVPA study of bilinguals, using Spanish and Basque translational equivalents as stimuli, was used to investigate how the depth of word processing (shallow vs. deep) affects the decoding of the semantic category from multi-voxel patterns of activity in putative semantic areas of the brain. The depth of processing was varied by motivating shallow processing i.e. just reading in half of the trials and relatively deeper processing i.e. reading accompanied by thinking of meaning in the other half of trials. Using cross-language decoding, it was investigated how the depth of processing mediates cross-language generalization of the semantic category in well-known semantic areas of the brain.

In the third empirical chapter (**chapter 4**), some lexico-semantic factors including frequency and concreteness of words were considered. Specifically, an fMRI-based MVPA study of bilinguals, using an extensive set of animal/non-animal Spanish and Basque translational equivalents with varying levels of concreteness and frequency as stimuli, was used to investigate how lexico-semantic factors influence the cross-language generalization of the semantic category in canonical substrates of the semantic network. The fact that this study used a comparatively larger set of stimuli gives it an edge over the other two studies in term of the types of multivariate analysis that can be conducted and the general inferences that can be drawn.

Finally, the last chapter i.e. the general discussion (**chapter 5**), aims to provide an overview of the results, relating the findings to the existing literature. The generalizability and theoretical implications of the results are considered. Additionally, the strengths and weaknesses of the presented studies are mentioned and some future prospects discussed.

Chapter 2

How State of Visual Awareness Influences the Cross-Language Generalization of Semantic Representations

2.1 Introduction

Visual word processing entails activation of a number of different processes, these range from orthographic operations in low-level visual areas (posterior areas including left occipitotemporal region and superior temporal gyrus; [73, 74, 75, 76, 77]) to semantic processes in mainly high-level association areas (a left-lateralized network of 7 regions including dorsomedial prefrontal cortex, inferior frontal gyrus and ventromedial prefrontal cortex; see a meta-analysis [1]) of the brain. An important question in the domain of non-conscious processing is: to what extent can the high-level cognitive processes unfold in the absence of conscious awareness [78, 79, 80, 81]. Whereas studies demonstrating non-conscious processing at relatively low levels of analysis (e.g. orthographic) are widely replicable and well-established now [82, 83, 84, 85], most of the evidence implying non-conscious processing at higher levels (e.g. semantic) has been subject to many criticisms for reasons we discuss below (see also [86, 87]).

Ever since the seminal work by Marcel in 1980, non-conscious semantic processing has been investigated using visual masked priming paradigm. In a typical such experiment, participants engage in a lexical decision task, and non-conscious access to semantics is said to occur if they respond faster to targets preceded by a semantically-related unconscious prime (e.g. cat-dog) as compared to targets preceded by a semantically-unrelated unconscious prime (e.g. bag-dog). Initial studies using this paradigm were criticised on several grounds [88, 89, 90, 91, 92], the most prominent being the methodological shortcomings in how the threshold of prime awareness was established. In studies like [93] and [94] for example, this threshold was established "offline", using separate blocks of detection trials prior to the semantic judgement trials, while in those including [95, 96, 97] and [98], it was assessed after the semantic categorization task. These approaches do not assess sensitivity to the primes in an "online" manner at the time of prime-target presentation (see [99]; p. 18) and therefore are prone to either overestimation of awareness due to perceptual learning throughout the whole experiment [100] or underestimation due to post-experiment fatigue and loss of motivation ([101]; for a detailed review of such issues, see [102, 103], and [104]). Another important objection raised by [95] and [105] regarding other semantic priming studies including [106, 107, 108] and [109] is that the non-conscious semantic effects in such studies are explainable by a direct association mapping between the stimulus and the motor response (S-R mapping). Abrams and Greenwald argued that these experiments used the same set of words as primes and targets, and often with a strict response deadline, this enabled the brain to develop a shallow stimulus-response association that bypassed semantic analysis (but see [110] for a study that circumvents this issue using a masked priming paradigm with number words). Another study by [111] presented masked emotional words with target-mask delay varied between the range of 33 milliseconds (ms) and 100 ms. They presented emotionally negative (e.g. "pain") and neutral words (e.g. "color"), and collected response to the word naming task and a visibility rating (on a quasi-continuous visual scale) after each word presentation. They showed that emotional words enjoy a better access to consciousness as compared to neutral words which was intepreted as reflecting preferential non-conscious processing of emotional words. Although this study provided evidence for non-conscious semantic processing, emotional words were used which are known to be processed extraordinarily quickly and automatically [112, 113, 114, 115].

The goal of this study is to provide evidence of non-conscious semantic processing that circumvents the key issues noted above, most notably, the known difficulties in demonstrating the lack of awareness. First, we used a combination of moment-to-moment subjective reports of (un)awareness with signal detection measures and then analyzed the patterns of brain activity for words that observers rated as unaware and which critically were associated with null behavioural discrimination performance. This approach mitigates the concerns associated with the "offline" assessment of awareness which is standard in subliminal priming studies, even when objective measures of awareness based on signal detection theory are used. As noted above signal detection thresholds can vary across different testing sessions and thus any assessment of awareness must ideally occur concurrently with trials that will be used to demonstrate behavioral or neural evidence of unconscious information processing. In particular, here we sought to find out brain-based evidence that the semantic category of words was processed even though participants lack sensitivity to the relevant information. Accordingly, we used multivariate pattern analysis (MVPA) of functional MRI signals to decode the semantic category of the items. A similar approach has recently been taken by [116], however, this study only involved the distinction between non-words and words embedded in sentences.

Thus, the first question we ask is whether the brain can encode the meaning of neutral words of animal and non-animal categories in the absence of conscious awareness. Additionally, we also aim to investigate the extent to which these non-conscious semantic representations of words are common or shared between different languages. Two recent fMRI studies [71] and [72] showed that a decoder trained to classify the meaning of words in one language can predict with above-chance performance the meaning of words in the other language. Since they found shared patterns in well-known semantic areas of the brain, both studies claim to have pinned down language-independent semantic representations. This is an important line of research as it explores the existence of conceptual representations that are supposed to be more general and associated with language-free perceptual experience [117]. We argue that the key limitation of these studies is that the words were fully visible and participants were required to consciously think about the properties of the words. Therefore, it still remains to be seen whether such language-independent semantic representations can also emerge in the patterns of brain activity in the absence of conscious awareness and null behavioral sensitivity.

2.2 Materials and Methods

2.2.1 Participants

Twenty four early and proficient Spanish-Basque bilinguals (mean age 22.3 ± 3.0 years; 17 female) including fourteen with Spanish as L1 were scanned using MRI. All of them had a normal or corrected to normal vision, gave written informed consent prior to the experiment and were financially compensated with 20 euros for their participation. The experiment lasted for about one and a half hour. Three of the participants were excluded before fMRI-based MVPA analysis. Two for failure to submit the category response in more than 50% of the trials, and one for failure to use the visibility ratings properly. The experiment was approved by the BCBL Ethics Review Board and complied with the guidelines of the Helsinki Declaration.

The online platform used for the recruitment (`www.bcbl.eu/participa`) also required participants to fill different questionnaires aimed at gathering information related to language proficiency of both languages. The collected data showed that all participants had acquired both languages before the age of 6. The mean age of acquisition was found to be 0.52 for Spanish and 1.05 for Basque with no statistically significant difference ($t(21) = -1.07, p = 0.30$). When considering their reported performance in the two well-known tests of language proficiency i.e. LexTALE (was available for only 20 out of 21 participants) [118] and BEST [119], statistically significant differences (LexTALE: $t(20) = 2.94; p < 0.05$, BEST: $t(21) = 5.15; p < 0.05$) were found between Spanish (LexTALE: 93.75 ± 4.62, BEST: 99.54 ± 1.13) and Basque (LexTALE: 87.22 ± 7.23, BEST: 86.46 ± 10.86). These scores thus show participants to be more proficient in Spanish than in Basque. Basque and Spanish are two very different languages with different roots. While Spanish is a romance language, Basque has unknown linguistic roots. It is an isolated pre-indo-european language. In addition, Basque holds many prominent linguistic differences with Spanish in the canonical word order in sentences regarding subject, verb and object, morphology (Basque: agglutinative), syntax (Basque: ergative), and lexicon (many different vocabulary and non-cognates).

2.2.2 MRI Acquisition

SIEMENS's Magnetom Prisma-fit scanner, with 3 Tesla magnet and 64-channel head coil, was used to collect, for each participant, one high-resolution T1-weighted structural image and eight functional images (corresponding to eight sessions). In each fMRI session, a multiband gradient-echo echo-planar imaging sequence with acceleration factor of 6, resolution of $2.4 \times 2.4 \times 2.4$ mm^3, TR of 850 ms, TE of 35 ms and bandwidth of 2582 Hz/Px was used to obtain 585 3D volumes of the whole brain (66 slices; FOV=210 mm). The visual stimuli was projected on an MRI-compatible out-of-bore screen using a projector placed in the room adjacent to the MRI-room. A small mirror, mounted on the head coil, reflected the screen for presentation to the participants. The head coil was also equipped with a microphone that enabled the participants to communicate with the experimenters in between the sessions.

2.2.3 Experimental Procedure

Each trial began with a fixation period of 500 ms followed by a blank screen of another 500 ms (see Figure 2.1). The target word, sandwiched between two 66 ms circular white noise masks, was presented for 66 ms and was followed by a response period of 3 s. During this period, the participants were asked two questions, one after another, and were supposed to respond to each within the respective time window of 1.5 s each. First, which semantic category does the word belong to, animals (A) or non-animals (nA)? To eliminate the effect of motor response difference on the choice of a semantic category, the mapping between choice and response button was randomly assigned on each trial. So, for some trials, A was on the right with nA on the left of the response screen, while for others, A was on the left with nA on the right. Participants were instructed to make their choice between left (i.e. button 1) and right (i.e. button 2) buttons based on the text displayed ("A nA" or "nA A") during the response period. Participants also provided an awareness rating of their visual experience of the word (1, 2 or 3); 1: I didn't see anything, 2: I think I saw a letter but not the word, or 3: I think I saw the word clearly or almost clearly. During training sessions, participants were given clear instructions that they were supposed to provide a forced-choice response to the category of the words in all the trials even for those in which they did not see the word at all (i.e. visibility rating of 1).

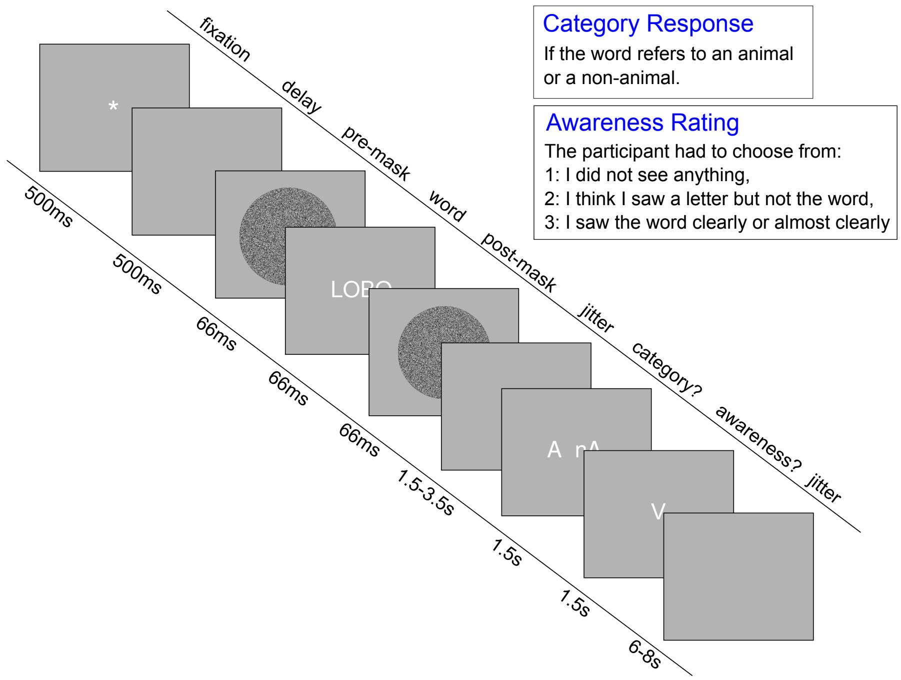


Figure 2.1: The figure summarizes the experimental design. A word was presented in the center of the screen for 66 ms. This was both preceded and followed by circular white noise masks with each lasting for 66 ms. Next, after a jittered interval of 1.5-3.5 s, participants responded to two questions: 1. Which category from among animals and non-animals does the word belong to? and 2. Which awareness rating from among 1, 2, and 3 does best describe his/her perceptual awareness of the word? The inter-trial interval that followed these responses was a jittered interval of 6-8s.

To ensure sufficient number of examples across the different states of visual awareness and to compensate for the changes in perceptual threshold across sessions [111], the luminance of the words was varied based on an adaptive staircase procedure. Specifically, this procedure increased the value of luminance by 0.02 if the participant pressed 1, decreased it by 0.01 if he/she pressed 2, and decreased it by 0.02 if the he/she pressed 3 for the awareness rating in the previous trial. The starting point of the first session's staircase was based on a pre-experiment calibration session; for subsequent sessions, the final luminance from the previous session was used. The pre-experiment calibration involved running two staircases, first with a luminance step size of 0.1 and then with 0.02 (just like the experiment), and were used to determine a threshold of luminance that consistently coincided with the detection failure rate of 40% (or 40% of trials being labeled as 1-rating: "I did not see anything").

A total of 8 words were used in the whole experiment. In all of the blocks of all of the sessions of the experiment, the same 8 words were presented either in Spanish or in Basque. These were 4 animal words including wolf, rooster, fox, sheep, and 4 non-animal words including candle, key, tube and mirror (for Spanish and Basque translations, see Figure 2.2). All these words were non-cognates and were balanced with respect to length and frequency (per million words; a standard measure independent of the corpus size) across categories (animals and non-animals) and across languages (see Table 2.1 for details) based on the statistics provided by Espal (for Spanish; [120]) and E-Hitz databases (for Basque; [121]). The requirement of length and frequency balancing across categories and languages put some constraints on the number of words, nevertheless the number finally selected was in keeping with previous studies of semantic decoding [71, 72, 122].

Both instructions and stimuli were presented at the center of the screen, in white against gray background and in all uppercase Arial font. The same stimuli was used for both the calibration and the actual experiment.

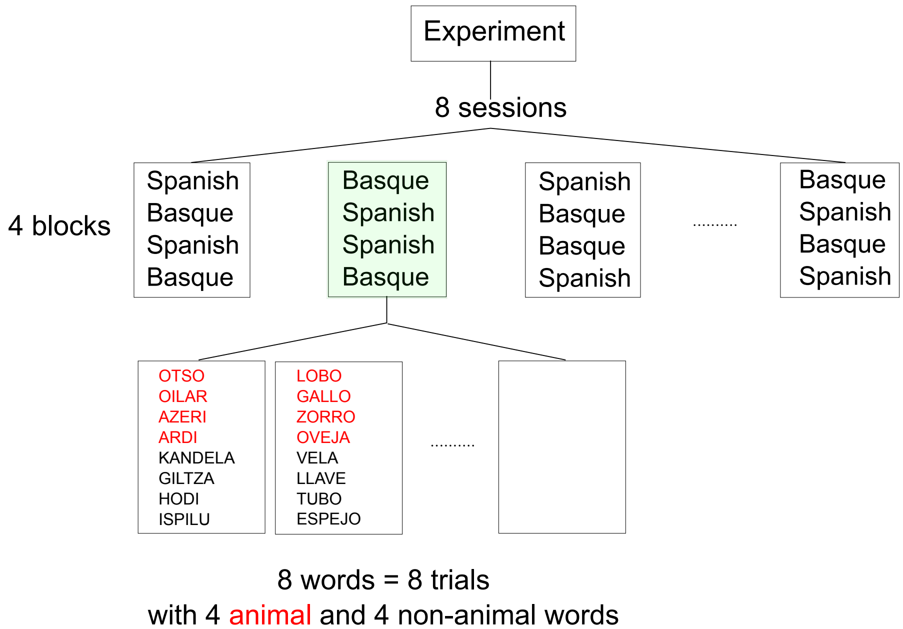


Figure 2.2: The figure summarizes the organization of sessions, blocks and trials in the experiment. Each experiment comprised of 8 sessions where each session was further subdivided into 4 language blocks (2 Spanish and 2 Basque). Each of these blocks was made up of 8 trials corresponding to single presentation of each of 4 animal and 4 non-animal Spanish/Basque words.

	Spanish		Basque	
	Animal	Non-animals	Animals	Non-animals
Length	4.5±0.58	4.75±0.96	4.5±0.58	5.25±0.96
Frequency	28.73±19.90	19.90±6.12	23.53±17.90	24.55±8.01

Table 2.1: The table shows mean word length and frequency of stimuli i.e. 8 animal and non-animal words with respect to both languages and semantic categories. These statistics were gathered using Espal for Spanish and E-Hitz for Basque. It can be seen that they were balanced across languages and categories.

The experiment was programmed and presented using Psychopy [123] and is summarized in Figure 2.2. Each fMRI session was subdivided into four language blocks (see Figure 2.1) with two Spanish (S) and two Basque (B) blocks, the order of these blocks was counterbalanced across sessions (SBSB, BSBS, and so on). In each of these blocks, eight words were presented (without repetition) in a random arrangement resulting in a total of thirty two trials per session.

To maximize the separation between the brain activity corresponding to stimuli and that related to response, the interval between post-mask and response was jittered between 1.5 s and 3.5 s. Similarly, to further facilitate the estimation of HRF, the inter-trial interval (ITI) was also jittered between 6 and 8 s. Both of these jitters were based on pseudo-exponential distributions resulting in 50% of trials with the ITI of 6 s, 25% with 6.5 s, 12.5% with 7 s and so on.

2.2.4 MRI Data Preprocessing

The preprocessing of fMRI data was performed using FEAT (fMRI Expert Analysis Tool), a tool in FSL suite (FMRIB Software Library; v5.0). After converting all data from DICOM to NIfTI format using MRIConvert (`http://lcni.uoregon.edu/downloads/mriconvert`), the following steps were performed on each session's fMRI. To ensure steady state magnetisation, the first 9 volumes corresponding to the task instruction period were discarded; to remove non-brain tissue, FSL's brain extraction tool (BET) [124] was used; head-motion was accounted for using MCFLIRT [125]; minimal spatial smoothing was performed using a gaussian kernel with FWHM of 3 mm and a high-pass filter with a cutoff of 90 s (calculated by FEAT's "Estimate High Pass Filter Tool" based on the analysis of the frequency content of the design). The sessions were coaligned by aligning each session to a reference volume of the already preprocessed first session. Further analysis was performed in native BOLD space. However, to be able to transform the anatomical region of interest (ROI) masks generated using Freesurfer (see below for details), transformation matrices were obtained using linear registration of BOLD scans to the structural space (and vice versa) based on 7 DoF global rescale transformation.

A set of 7 left-lateralized ROIs were pre-specified (see Figure 2.3) based on a meta-analysis of the semantic system carried out by [1]. This meta-analysis is most relevant because it identifies the most critical semantic areas using only fMRI studies that used words as stimuli. These identified ROIs include: inferior parietal lobe (IPL), lateral temporal lobe (LTL), ventromedial temporal lobe (VTL) including fusiform gyrus and parahippocampal gyrus, dorsomedial prefrontal cortex (dmPFC), inferior frontal gyrus (IFG), ventromedial prefrontal cortex (vmPFC), and posterior cingulate gyrus (PCG) along with precuneus. First, automatic segmentation of the high-resolution structural image was obtained using FreeSurfer's automated algorithm `recon-all`. Next, `mri_binarize` was used to extract individual gray matter masks from aparc+aseg volume using corresponding label indices in FreeSurferColorLUT text file (https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/AnatomicalROI). And finally, after visually inspecting these in FSLView, they were transformed to each session's functional space using FLIRT [125, 126].

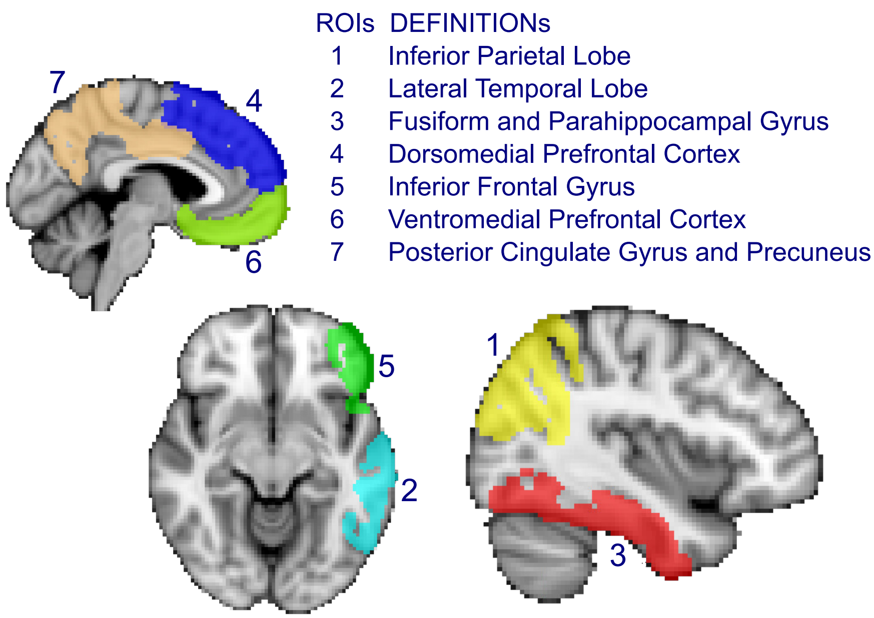


Figure 2.3: The figure shows the selected regions of interest projected on an MNI standard template image. These left-lateralized areas were pre-specified based on a meta-analysis by [1] and included inferior parietal lobe (IPL), lateral temporal lobe (LTL), ventromedial temporal lobe (VTL) including fusiform gyrus and parahippocampal gyrus, dorsomedial prefrontal cortex (dmPFC), inferior frontal gyrus (IFG), ventromedial prefrontal cortex (vmPFC), and posterior cingulate gyrus (PCG) along with precuneus.

2.2.5 Multivariate Pattern Analysis

Multivariate pattern analysis were conducted using scikit-learn [127] and Py MVPA [128] libraries. Specifically, classification based on a supervised machine learning algorithm i.e. linear support vector machine [129], was used to evaluate whether multi-voxel patterns in each of the seven ROIs carry information related to the semantic category (animal, non-animal) of the word in each state of awareness. Within-language (or language-specific) decoding involved restricting the analysis to trials of a specific language (either Spanish or Basque) while cross-language decoding entailed training the classifier on trials from one language and testing it on trials from another language. Both of these analysis were done separately for each of the awareness conditions. Additional details related to the data preparation, feature selection, classification and statistics are presented in the following subsections.

2.2.5.1 Data Preparation

For each subject, the relevant time points or scans of the preprocessed fMRI data of each session were labeled with attributes such as category and language using a Python script with corresponding Psychopy generated data files as input. The trial-by-trial awareness reports were used to separate the trials into 1-rating, 2-rating, and 3-rating trials. Invariant features were removed. These were the voxels/features whose value did not vary throughout the length of one session. If not removed, such features can cause numerical difficulties with procedures like z-scoring of features. Next, data from all eight sessions was stacked and each voxel's data points were session-wise z-score normalized and linear detrended. Finally, to account for the hemodynamic lag, one example was created per trial by averaging the 4 volumes between the interval of 3.4 s and 6.8 s after the word onset. Since the visibility rating of 1 represented the awareness report "I didn't see anything" and the mean behavioral performance in the corresponding 1-rating trials was also found to be at chance-level, these trials were considered as non-conscious trials. Similarly, trials with rating of 3 ("I think I saw the word clearly or almost clearly") were labeled as conscious trials. However, due to some participants having only a small number of 2-rating and others having a small number of 3-rating trials (see 2.3.2), both 2-rating and 3-rating trials were collapsed and were considered to represent one condition. It is worth noting however that the rating of 2 ("I think I saw a letter but not the word") does not represent a conscious state. Hence, the resulting combination of both 2-rating and 3-rating conditions were labeled as partially conscious.

2.2.5.2 Pattern Classification

Linear support vector machine (SVM) classifier, with all parameters set to default values as provided by the scikit-learn package ($l2$ regularization, $C = 1.0$, $tolerance = 0.0001$), was used for both within-language decoding and cross-language decoding in both partially conscious and non-conscious. The following procedure was repeated for each ROI separately. To obtain an unbiased generalization estimate, following [130], the data was randomly shuffled and resampled multiple times to create 300 sets of balanced train-test (80%-20%) splits. Since each example was represented by a single feature vector with each feature a mean of voxel intensities across the sub-interval of 3.4 s and 6.8 s (see 2.2.5.1), the length of a vector was equal to the number of voxels in the ROI. To further reduce the dimensionality of the data and thus reduce the chances of overfitting [131, 132], Principal Component Analysis (PCA) with all parameters set to default values as provided by the scikit-learn package (see `https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html`) was used. The number of components was equal to the number of examples thus resulting in all ROIs having equal number of components. These components were linear combinations of the preprocessed voxel data and since none of the components was excluded, it was an information loss-less change of the coordinate system to a subspace spanned by the examples [133]. Features thus created were used to train the decoder, and the classification performance on the test set was recorded. This procedure was repeated separately for each of the 300 sets, and the mean of corresponding accuracies was collected and averaged for each of the participants.

2.2.5.3 Statistics

To determine whether the observed decoding accuracy in an ROI is statistically significantly different from the chance-level of 0.5 (or 50%), a two-tailed t-test was performed with p-values corresponding to each of the ROIs corrected for multiple comparisons using a false discovery rate (FDR) method. To get the empirical estimate of chance-level, we ran the classification tests while randomly permuting over the category labels. The chance-level was computed across participants, ROIs, classification problems (within and cross-language) and states of awareness. For each case, 300 permutations were performed and the mean and standard deviation of the collected permutation scores was calculated across participants. For all ROIs, and classification problems, the chance-level was consistently found to be centered around 0.5. All effect sizes are reported as *mean effect size*±*standard error*; *t(degrees of freedom)=t-value; p-value* across all participants.

2.3 Results

2.3.1 Behavioral Results

2.3.1.1 Awareness ratings were used properly

To establish whether the word in each trial was consciously perceived or not, participants were asked to submit both the objective categorization response (animal or non-animal) and the subjective visibility response (on the scale of 1 to 3; see 2.2.3 for corresponding definitions) after each word presentation. Based on these responses, more than 40% of trials were found to be non-conscious (1-rating) in both Spanish (41 ± 4%) and Basque (45 ± 3%). Importantly, considering the objective performance in the animal vs. non-animal discrimination on these non-conscious trials, it was found to be at chance-level in both Spanish ($mean = 51 \pm 9\%; t(21) = 0.73; p = 0.47$) and Basque ($mean = 53 \pm 10\%; t(21) = 1.36; p = 0.18$; see Figure 2.4). Figure 2.4 also shows the objective performance for partially conscious trials. Specifically, it was found to be above chance for both 2-rating ($78 \pm 13\%; t(21) = 9.30; p < 0.05$ for Spanish and $74 \pm 14\%; t(21) = 7.67; p < 0.05$ for Basque), and 3-rating trials ($97 \pm 3\%; t(21) = 66.61; p < 0.05$ for Spanish and $97 \pm 4\%; t(21) = 49.99; p < 0.05$ for Basque). Taken together, this suggests that the participants used the awareness ratings correctly and the trials judged as not visible were genuinely invisible as per both subjective and objective behavioral measures.

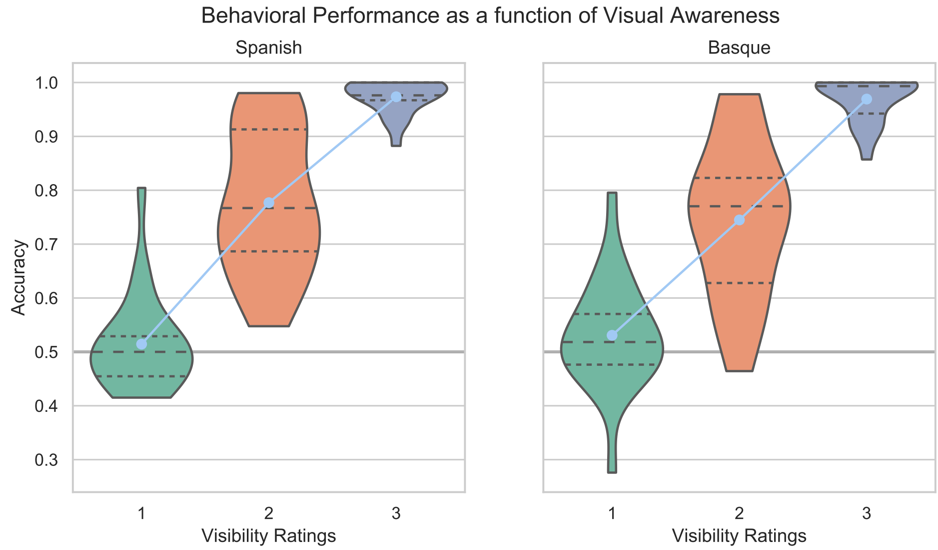


Figure 2.4: In trials with visibility rating of 1, the objective categorization performance was found to be at chance-level in both Spanish (left) and Basque (right). In those with visibility rating of 2, and 3, it was found to be clearly above chance in both the languages. The three dotted lines inside each violin are the quartiles. The black horizontal line in the background indicates the chance-level performance and the blue line shows the trend followed by the mean performance.

2.3.1.2 Stimulus strength was identical in all conditions

To compensate for variation in perceptual threshold across the experimental sessions, we decided to keep the adaptive luminance staircase (for details, see 2.2.3) running throughout the whole experiment. The average luminance of the words however was found to be similar across the different visibility conditions (see Figure 2.5). One way ANOVA with three levels showed no statistically significant difference between conditions ($F(21) = 0.50; p = 0.61$ for Spanish, and $F(21) = 0.50; p = 0.66$ for Basque) and therefore can be used to conclude that the stimulus strength was similar in both partially conscious and non-conscious conditions.

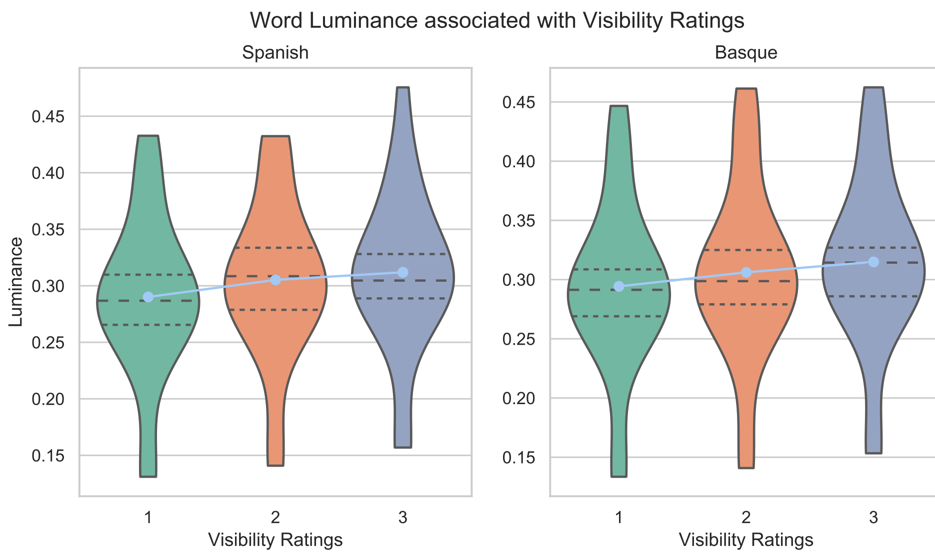


Figure 2.5: The figure shows the distribution of luminance values corresponding to each of the visibility ratings for both Spanish (left) and Basque (right). One way ANOVA with three levels showed no statistically significant difference between conditions. The three dotted lines inside each violin are the quartiles. The blue line passes through the means of the three distributions.

2.3.2 Brain Imaging Results

The primary goal of this study was to investigate whether the semantic category of non-conscious words can be predicted from BOLD response patterns, and which brain areas of the semantic network were involved. This classification problem was conducted separately for each language (henceforth within-language decoding). The second goal of the study was concerned with cross-language generalization, namely, whether it is possible to decode the meaning of the non-conscious words in one language using a decoder trained to do the same in another language? [71, 72].

Decoding was conducted separately for each of the awareness states. The average number of trials per subject was 82 for 3-rating (44 for Spanish, 38 for Basque), 65 for 2-rating (32 for Spanish, 33 for Basque), and 113 for 1-rating (54 for Spanish, 59 for Basque). Recall that since participants reported having no conscious awareness whatsoever in 1-rating trials and corresponding discrimination performance was also found to be at chance-level, these trials were considered as non-conscious trials. On the same lines, 3-rating trials (defined as "I saw a word clearly or almost clearly") reasonably qualified to be considered as conscious trials. However, since some participants had a small number of 3-rating trials ("I saw a word clearly or almost clearly"; $mean = 34\%; SD = 7\%$ in Spanish and $29\% \pm 7\%$ in Basque) and the others had a small number of 2-rating trials ("I think I saw a letter but not the word"; $mean = 25\%; SD = 10\%$ in Spanish and $26\% \pm 9\%$ in Basque), it was decided to combine both 2-rating and 3-rating trials and consider them as partially conscious trials (see 2.2.5.1). The high variability in 2-rating and 3-rating trials was likely due to constraints imposed by our paradigm, namely the adaptive staircase procedure biased the luminance of the words to maximize the number of non-conscious trials. So, it is only after combining both 2-rating, and 3-rating trials that the mean and variability became comparable to that of 1-rating trials and we obtained a more reasonable number for decoding.

2.3.2.1 Within-language Decoding

Within-language decoding involved restricting the SVM-based classification analysis to one language at a time, Figures 2.6 and 2.7 therefore present summary statistics of the ROIs for partially conscious and non-conscious conditions in both Spanish and Basque respectively. It can be seen that in non-conscious trials, considering Spanish results, the classification of the semantic category (animal/non-animal) was found to be statistically significantly above-chance in four out of seven ROIs including IPL ($mean = 54.7 \pm 7.0\%; t(21) = 2.97$; corrected $p = 0.02$; all p-values hereafter are FDR corrected), LTL ($53.0 \pm 7.4\%; t(21) = 1.82; p = 0.10$), VTL ($52.7 \pm 7.7\%; t(21) = 1.55; p = 0.14$), dmPFC ($56.1 \pm 6.4\%; t(21) = 4.25; p = 0.003$), IFG ($53.7 \pm 5.6\%; t(21) = 3.02; p = 0.02$), vmPFC ($53.9 \pm 7.6\%; t(21) = 2.27; p = 0.05$), and PCG ($54.0 \pm 7.5\%; t(21) = 2.42; p = 0.04$). In Basque, it was found to be above-chance in two of the seven ROIs including IPL ($mean = 52.8 \pm 6.5\%; t(21) = 1.94; p = 0.12$), LTL ($51.4 \pm 6.3\%; t(21) = 1.00; p = 0.38$), VTL ($54.4 \pm 6.5\%; t(21) = 3.04; p = 0.02$), dmPFC ($53.0 \pm 6.2\%; t(21) = 2.19; p = 0.09$), IFG ($51.8 \pm 6.0\%; t(21) = 1.34; p = 0.27$), vmPFC ($50.8 \pm 7.2\%; t(21) = 0.50; p = 0.07$), and PCG ($54.7 \pm 6.0\%; t(21) = 3.51; p = 0.02$).

In partially conscious trials, the classification of the semantic category was found to be statistically significantly above chance in all ROIs in both Spanish and Basque. Notably, while the decoding accuracies were similar in magnitude to that in non-conscious condition, above-chance accuracies were found to be distributed across all ROIs. For Spanish, these were: IPL ($mean = 54.1 \pm 4.5\%; t(21) = 4.08; p = 0.001$), LTL ($52.6 \pm 4.9\%; t(21) = 2.37; p = 0.03$), VTL ($52.9 \pm 4.4\%; t(21) = 2.93; p = 0.01$), dmPFC ($53.5 \pm 5.7\%; t(21) = 2.71; p = 0.02$), IFG ($55.9 \pm 4.8\%; t(21) = 5.51; p = 0.0002$), vmPFC ($52.3 \pm 4.7\%; t(21) = 2.24; p = 0.037$), and PCG ($55.7 \pm 5.7\%; t(21) = 4.51; p = 0.0008$). And for Basque, these were: IPL ($52.7 \pm 5.8\%; t(21) = 2.10; p = 0.049$), LTL ($54.8 \pm 4.7\%; t(21) = 4.59; p = 0.001$), VTL ($54.4 \pm 5.0\%; t(21) = 4.00; p = 0.002$), dmPFC ($52.8 \pm 5.4\%; t(21) = 2.28; p = 0.039$), IFG ($54.1 \pm 6.8\%; t(21) = 2.75; p = 0.02$), vmPFC ($52.9 \pm 4.7\%; t(21) = 2.66; p = 0.02$), and PCG ($53.9 \pm 6.3\%; t(21) = 2.80; p = 0.02$).

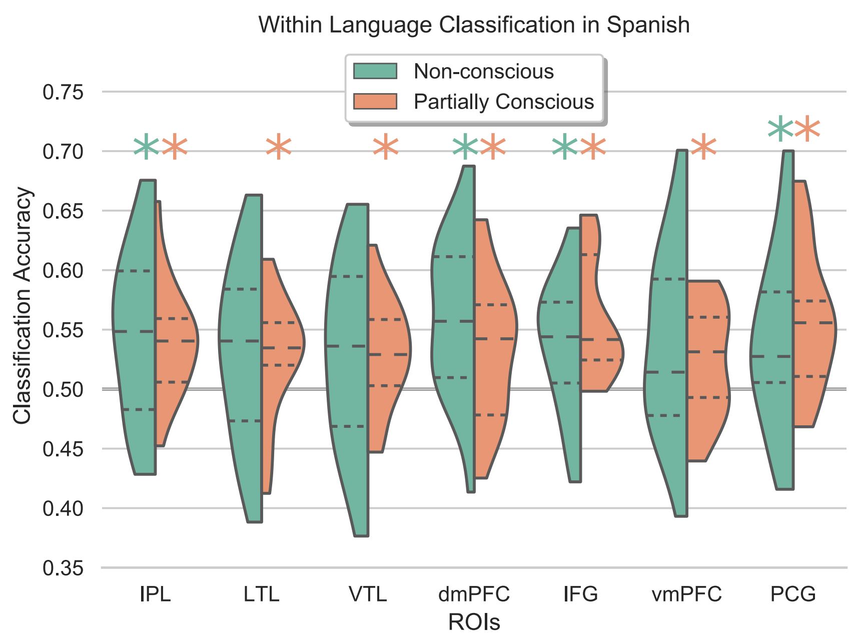


Figure 2.6: The figure shows the summary statistics of the ROIs for both partially conscious and non-conscious within-language decoding in Spanish. It can be seen that the decoding was above chance in four ROIs in non-conscious but all seven ROIs in partially conscious condition. The three dotted lines inside each violin are the quartiles. Orange and green asterisks signify statistically significantly above chance decoding in partially conscious and non-conscious conditions respectively. The black horizontal line in the background indicates the chance-level performance.

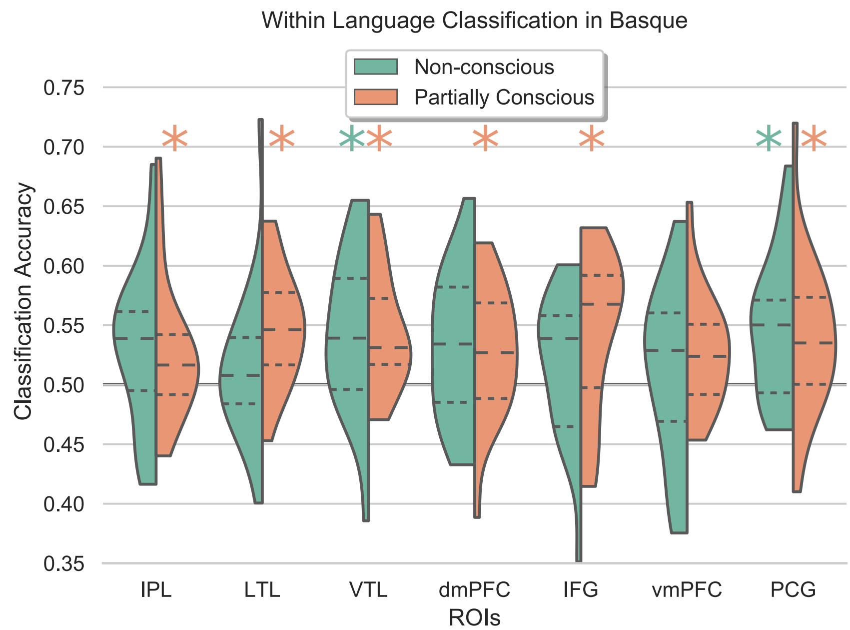


Figure 2.7: The figure shows the summary statistics of the ROIs for both partially conscious and non-conscious within-language decoding in Basque. It can be seen that the decoding was above chance in two ROIs in non-conscious but all seven ROIs in partially conscious condition. The three dotted lines inside each violin are the quartiles. Orange and green asterisks signify statistically significantly above chance decoding in partially conscious and non-conscious conditions respectively. The black horizontal line in the background indicates the chance-level performance.

It is noteworthy that there was one subject in Spanish and another in Basque whose behavioral discrimination performance was found to be around 80% in non-conscious condition. This is reminiscent of a "blindsight" effect or perception without awareness [134, 135]. These represented outliers because their behavioral performance was 3 standard deviations higher than the mean. Although this was related to the behavioural performance, we wanted to ensure that these participants were not driving the above-chance decoding in non-conscious (see Figures 2.6 and 2.7). Therefore, within-language decoding procedure was re-run without including these outlier participants. Notably, it was found that that the pattern of results in the non-conscious remains intact in both Spanish (see Figure 2.8a) and Basque (Figure 2.8b). Specifically, in Spanish (see Figure 2.8a), the decoding of the semantic category (animal/non-animal) was again found to be statistically significantly above-chance in IPL ($mean = 55.0 \pm 7.0\%; t(21) = 3.08; p = 0.01$), dmPFC ($56.0 \pm 6.6\%; t(21) = 3.98; p = 0.006$), IFG ($54.0 \pm 5.6\%; t(21) = 3.08; p = 0.01$), and PCG ($54.7 \pm 7.1\%; t(21) = 2.86; p = 0.02$). Similarly in Basque (see Figure 2.8b), it was found to be above-chance in two of the seven ROIs i.e. VTL ($54.7 \pm 6.5\%; t(21) = 3.13; p = 0.02$) and PCG ($54.6 \pm 6.1\%; t(21) = 3.51; p = 0.02$).

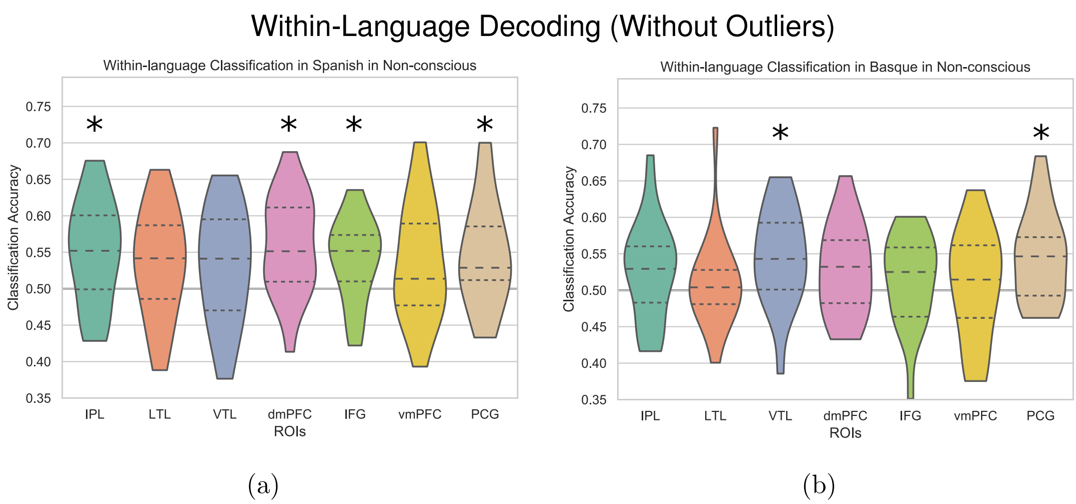
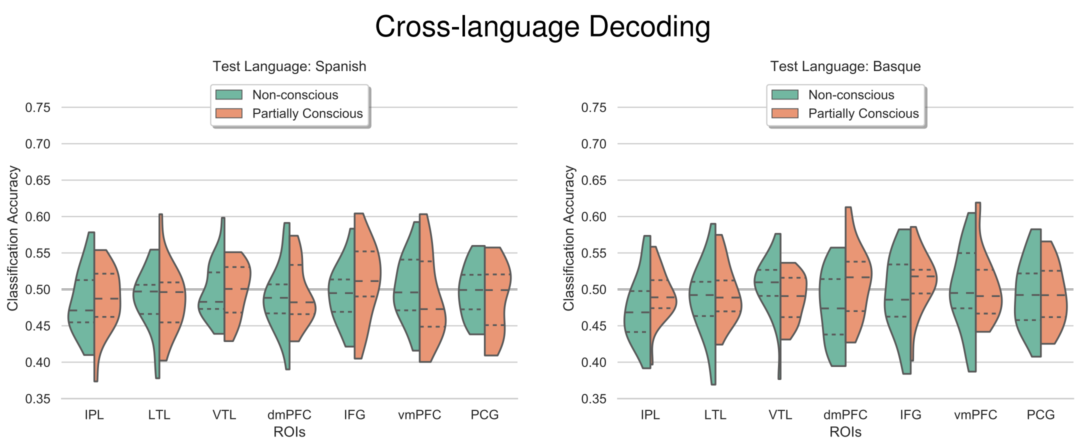


Figure 2.8: There was one subject in Spanish and another in Basque whose behavioral discrimination performance was found to be around 80% in non-conscious condition. The figures show that the pattern of results in non-conscious decoding remained intact even after the removal of these outlier participants. Specifically, the same ROIs were found to be critical for the decoding of meaning in both Spanish (Figure 2.8a) and Basque (Figure 2.8b). The black asterisk signifies statistically significantly above chance decoding.

2.3.2.2 Cross-language Decoding

Cross-language decoding involved training the decoder on the examples of one language (train language) and testing it on the examples of the other language (test language). So, with Spanish as test language, Basque was the train language and vice versa. Figure 9 presents summary statistics of the ROIs for both partially conscious (for generalization from Spanish to Basque: IPL ($mean = 49.1 \pm 3.4\%; t(21) = -1.18; p = 0.49$), LTL ($49.4 \pm 3.8\%; t(21) = -0.69; p = 0.58$), VTL ($48.9 \pm 3.0\%; t(21) = -1.66; p = 0.49$), dmPFC ($51.0 \pm 4.8\%; t(21) = 0.96; p = 0.49$), IFG ($51.1 \pm 3.9\%; t(21) = 1.22; p = 0.49$), vmPFC ($50.3 \pm 4.6\%; t(21) = 0.30; p = 0.77$), and PCG ($49.1 \pm 4.1\%; t(21) = -1.04; p = 0.49$); for generalization from Basque to Spanish: IPL ($mean = 48.5 \pm 4.2\%; t(21) = -1.53; p = 0.51$), LTL ($48.6 \pm 4.5\%; t(21) = -1.43; p = 0.51$), VTL ($50.0 \pm 3.6\%; t(21) = -0.03; p = 0.98$), dmPFC ($49.5 \pm 4.1\%; t(21) = -0.57; p = 0.67$), IFG ($51.1 \pm 5.2\%; t(21) = 0.97; p = 0.51$), vmPFC ($48.8 \pm 5.6\%; t(21) = -0.93; p = 0.51$), and PCG ($48.8 \pm 4.6\%; t(21) = -1.13; p = 0.51$)) and non-conscious conditions (for generalization from Basque to Spanish: IPL ($mean = 48.3 \pm 4.4\%; t(21) = -1.76; p = 0.65$), LTL ($48.7 \pm 4.3\%; t(21) = -1.35; p = 0.67$), VTL ($49.8 \pm 3.8\%; t(21) = -0.22; p = 0.92$), dmPFC ($48.9 \pm 4.6\%; t(21) = -1.02; p = 0.75$), IFG ($49.4 \pm 4.0\%; t(21) = -0.70; p = 0.86$), vmPFC ($50.1 \pm 4.4\%; t(21) = 0.11; p = 0.92$), and PCG ($49.7 \pm 3.7\%; t(21) = -0.41; p = 0.92$); for generalization from Basque to Spanish: IPL ($mean = 47.4 \pm 4.6\%; t(21) = -2.53; p = 0.14$), LTL ($48.6 \pm 5.0\%; t(21) = -1.23; p = 0.54$), VTL ($50.4 \pm 4.1\%; t(21) = 0.40; p = 0.81$), dmPFC ($47.8 \pm 4.9\%; t(21) = -2.01; p = 0.20$), IFG ($49.2 \pm 5.4\%; t(21) = -0.70; p = 0.81$), vmPFC ($49.8 \pm 5.6\%; t(21) = -0.16; p = 0.88$), and PCG ($49.4 \pm 4.7\%; t(21) = -0.56; p = 0.81$)) and it can be seen that in both conditions, the cross-language generalization was at chance-level in all ROIs.



(a)

(b)

Figure 2.9: The figures show the summary statistics of the ROIs for both partially conscious and non-conscious cross-language decoding. Specifically, the Figure 2.9a corresponds to when the decoder was trained on Basque and tested on Spanish, and the Figure 2.9b is for when the decoder was trained on Spanish and tested on Basque. From both of these figures, it is clear that the cross-language generalization was at chance-level in both partially conscious and non-conscious conditions. The three dotted lines inside each violin are the quartiles. The black horizontal line in the background indicates the chance-level performance.

In a further test of cross-language generalization, we combined the data from all ROIs in order to potentially increase the chances of decoding. However we found chance-level cross-language generalization in both partially conscious ($mean = 49.7 \pm 3.5\%; p = 0.69$ for Spanish and $49.5 \pm 4.2\%; p = 0.57$ for Basque) and non-conscious conditions ($48.8 \pm 5.6\%; p = 0.36$ for Spanish and $mean = 49.0 \pm 4.5\%; p = 0.34$ for Basque). Within-language decoding was however statistically above chance-level in both partially conscious ($54.4 \pm 5.4\%; p = 0.001$ for Spanish and $mean = 55.2 \pm 5.5\%; p = 0.0004$ for Basque) and non-conscious conditions ($53.2 \pm 6.6\%; p = 0.04$ for Spanish and $mean = 55.0 \pm 6.9\%; p = 0.004$ for Basque) similar to the results found before.

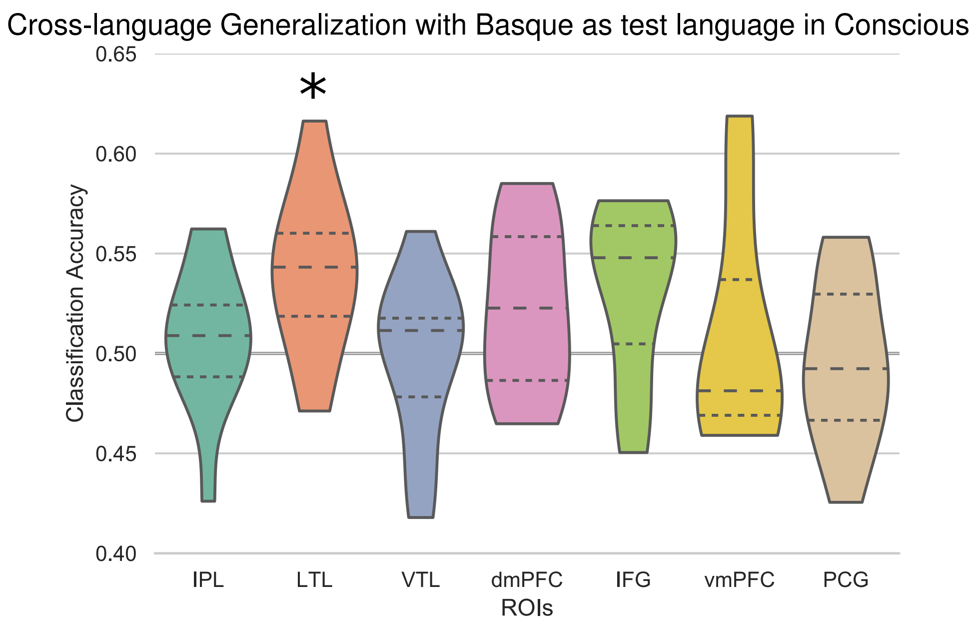


Figure 2.10: The figure shows the summary statistics of the ROIs for cross-language generalization from Spanish to Basque. Only participants with relatively high within-language decoding performance were included. It can be seen that one ROI showed statistically significantly above-chance generalization from Spanish to Basque. The three dotted lines inside each violin are the quartiles. The black horizontal line in the background indicates the chance-level performance.

It could be argued that the absence of cross-language decoding could be due to a floor effect, namely, given that classification accuracy was just above chance in the within-language decoding, it could only drop to chance level in the cross-language generalization. To mitigate the presence of floor effects that could affect the ability to find cross-language generalization, we ran an analysis including only those participants that had relatively high within-language classification performance (i.e. greater than or equal to 60% in the unconscious condition). There were 12 participants that satisfied this criteria for Spanish, and 7 that did it for Basque. Notably, in fully conscious condition, we found that there were two ROIs that showed statistically significantly above-chance cross-language generalization from Spanish to Basque i.e. LTL ($54.0 \pm 3.9\%; p = 0.006; t(12) = 3.44$) and IFG ($53.1 \pm 4.1\%, p = 0.028, t(12) = 2.52$), with the LTL surviving the correction for multiple comparisons (see Figure 2.10). The generalization from Basque to Spanish ($N = 7$) was found to be at chance-level in all ROIs. We did not find any evidence of cross-language generalization in partially conscious and non-conscious trials.

Finally, we also addressed decoding accuracy on fully conscious 3-rating trials. Because of the constraints in the number of 3-rating trials across participants (see 2.3.2), we looked for those that had at least 20 animal and 20 non-animal examples with the visibility rating of 3. There were 11 participants that satisfied this criteria for Spanish and 5 that did it for Basque. In within-language, we did not find significant gain in performance as compared to corresponding partially conscious and non-conscious results. Cross-language generalization was again found to be at chance-level in all ROIs.

2.4 Discussion

Our study investigated the brain basis of non-conscious semantic processing using masked word paradigm. Using multivariate pattern analysis of BOLD responses, we provide new insight into the brain substrates that support semantic representations across distinct states of visual awareness. Specifically, we showed that BOLD activity patterns associated with non-conscious words contain information that allows for decoding of the category of words both in Spanish and Basque (i.e. within language decoding). Notably, in the present study the words were non-conscious according to both subjective (i.e. rated as fully unaware on trial-by-trial basis [136] as well as objective measures given that behavioural discrimination of the word category was found to be at chance level.

ROI analysis (see 2.2.5.2) showed that such discriminative patterns for non-conscious items were found in canonical areas [1, 2] of the semantic network. Specifically, above-chance classification accuracies were found in a rather distributed set of brain regions including IPL, dmPFC, IFG and PCG for Spanish, and VTL and PCG for Basque. All of these areas have previously been associated with semantic processing of visible words in studies involving animals-tools and animals-artifacts contrasts [137, 138, 139, 140]. We also showed that for partially conscious trials of both Spanish and Basque, such discriminative BOLD patterns were even more distributed, namely, significant decoding was found in all pre-specified left-lateralized seven ROIs of the semantic network. On the other hand, addressing the second question i.e. cross-language generalization of semantic representations, we found little evidence for semantic generalization across languages, even on conscious trials.

All seven canonical areas of the semantic network were found to be implicated in the representation of word category under conditions where participants showed some awareness of the words in both Spanish and Basque. These results go in line with previous decoding studies of word meaning including [141, 142] and [122]. In the non-conscious condition, only one ROI was found to be common between Spanish and Basque i.e. PCG. Furthermore, while four ROIs (IPL, dmPFC, IFG, and PCG) were implicated in non-conscious semantic processing in Spanish, only two ROIs (VTL, PCG) were found for Basque. However, this pattern of results should not be taken to suggest that there are language-specific semantic representations. Different factors may have contributed to this pattern of results. For instance, our group of participants was mixed with most having Spanish as L1. Also, whereas no statistically significant difference was found between the age of acquisition of Spanish and Basque, the performance at both LexTALE and BEST tests of language proficiency was found to be statistically significantly superior in Spanish as compared to Basque (see 2.2.1). Finally, we compared the decoding accuracy in those ROIs that did not overlap between Spanish and Basque but no difference between languages was found. Taken together, it is possible that inter-individual variability may have promoted the absence of a complete overlap between the informative ROIs between Spanish and Basque.

In the non-conscious trials of Spanish (see Figure 2.6), besides in IPL, and

PCG, we also found significant decoding of meaning in dmPFC and IFG. What is interesting with the involvement of these frontal areas in non-conscious semantic representations is that it has implications for theoretical models of conscious and non-conscious processing i.e., Global Workspace theory [143]. According to this model, conscious representations result from widely distributed activity patterns involving both anterior (e.g. PFC) and posterior areas (e.g. object-selective brain areas), and information is broadcasted in these areas by means of top-down recurrent processing. Neuroimaging studies using masked priming paradigms indicate that non-conscious orthographic processing of words can occur in the left fusiform gyrus (i.e. the visual word form area; [144, 145]). Priming experiments indicate that the non-conscious semantic priming implicates the left superior temporal areas [94]. Additional results from event related potentials indicate non-conscious semantic processing indexed by the N400 [146, 147, 148] but see [149]. Although these studies can be criticized based on the issues highlighted in the introduction (i.e. the absence of trial-by-trial measures of awareness, see 2.1), the pattern of results suggests a relatively localized regional activity in non-conscious word processing that does not implicate higher-level prefrontal areas typically associated with conscious semantic processing (i.e. the left inferior frontal cortex) [1]. The present results, on the other hand, indicate that the non-conscious semantic representations can be encoded in relatively distributed brain substrates involving the prefrontal cortex. A key difference between our paradigm and masked priming paradigm is that here the words were task-relevant and in masked priming, the primes are task-irrelevant. There is a limited data that supports the involvement of frontal areas during non-conscious word processing. One prior masked priming study [150] showed regional BOLD response changes in a left-lateralized set of brain regions including the inferior frontal gyrus, inferior parietal and lateral temporal lobes during non-conscious processing of masked words. Another [116] recently showed that the meaningful sentences rendered non-conscious by continuous flash suppression could be discriminated from non-words by using fMRI-based MVPA, specifically in left-lateralized brain areas including superior temporal sulcus and the middle frontal gyrus. However, our study goes beyond this finding, and shows that not only lower-level structural representations can be isolated [116], but the semantic category of non-conscious words can also be classified. The present results also align with the prior research in visual working memory and executive control, which also indicates that dorsolateral prefrontal regions can be implicated in processing and brief maintenance of non-conscious visual stimuli ([81, 151, 152, 153]; though prefrontal activity in this later study occurred for subjectively unaware items unlike for items associated with null behavioral discrimination as demonstrated here). However, it is likely that non-conscious representations in prefrontal cortex are weak and hence unlikely to ignite sustained and strong feedback processing loops in distributed brain networks, which can be a requirement for information to become conscious [154]. Further research is needed to understand the limits and the functional scope of non-conscious semantic representations in the human brain, for instance, by testing its durability and the temporal dynamics of distributed semantic networks.

We now turn to the cross-language generalization results. All of the ROIs showed chance-level decoding accuracy for semantic generalization from Spanish to Basque and vice versa. This happened not just for non-conscious words but also for partially conscious trials. We only found some evidence of across language generalization from Spanish to Basque in the conscious trials when we restricted our analysis to those participants with within-language decoding accuracies well exceeding chance level (i.e. 0.6), in order to avoid the presence of floor effects in cross-language generalization.

This is the first time that MVPA-based cross-language generalization has been used to investigate the scope of non-conscious semantic representations. However, the same approach has already been used with positive results in a number of different fMRI studies where words were available to conscious awareness [71, 72, 155, 156]. The factors leading to cross-language generalization are not well understood. There are a number of reasons that can explain why we did not find strong evidence for it. Firstly, the experiment was designed to maximize the number of non-conscious trials. The stimuli was briefly presented and masked, and luminance varied based on a staircase procedure that was biased towards decreasing luminance in response to ratings of partial or full awareness. Therefore, even though the participants reported partial and full visibility of the items, this does not mean that the stimuli strength was comparable to that of previous studies that reported cross-language generalization [71, 72], where stimuli were presented for much longer durations, were fully conscious and even observers were asked to think about the items to ensure that deep semantic analysis is taking place. Accordingly, our task may only have promoted shallow encoding of the words. Given the relatively small number of words used, it is also possible that the observers learned a mapping between the properties of the word stimuli and the semantic categorization response, which did not involve the level of processing required for across language generalization. We suggest that our task may have promoted a level of processing that is sufficient for within-language decoding but insufficient for cross-language generalization.

It is also worth noting here that a significant amount of behavioural studies have addressed language-independent semantic representations by using translation and associative masked priming. Notably, while some of these studies have succeeded at showing cross-language semantic priming, most of them suffer from a number of methodological issues [157, 158]. For instance, the reliance on post-hoc assessment of the visibility of prime words (and the absence of trial-by-trial measures of awareness) make it hard to establish that priming effects are not contaminated by some trials with prime awareness [102, 103, 104]. Further, the use of long SOAs (**s**timulus **o**nset **a**synchrony i.e. the time for which the prime gets displayed before it gets replaced by a target) do not rule out the operation of conscious strategic processes. Notably, non-replicable findings have been observed with most studies reporting absence of effect [45, 158] to a few reporting statistically significant cross-language facilitation [44], yet trial-by-trial awareness assessment was not used in this study either. It is probably in the light of these issues that [158] go so far as to conclude that all the cross-language priming effect seems to be the result of an improper control of additional conscious strategic factors that result in significant cross-language facilitation.

The current study demonstrated that the meaning of non-conscious words can be encoded in multi-voxel patterns of activity in putative semantic regions, including frontal areas. Whereas within-language classification of word meaning is possible in non-conscious contexts, cross-language generalization (or evidence for language-independent semantic representations) seems harder to isolate; the latter may require not just conscious perception but a deeper semantic analysis too. Additional work is needed to make this determination.

Chapter 3

The Effect of Depth of Processing on the Brain Representation of Meaning Across Languages

3.1 Introduction

A key unresolved question is whether different languages in bilingual people are integrated in the same system with shared/overlapping representations or rely on separate systems/representations for each language. Behavioral evidence from cross-language priming studies suggests that semantic representations are at least partially overlapping [43, 44, 45]. The evidence has led to the development of psycholinguistic models of bilingual language representation [46, 47]. Although these models differ in their predictions about the mechanisms that underlie lexical processing and the links between lexical and semantic processing of the two languages, they agree that semantic representations are at least partially overlapping between languages. Yet, other studies have failed to support overlapping semantic systems [48, 49, 50]. The mixed evidence between the cross-language priming studies is likely to originate from a lack of control of low-level properties of primes and targets (e.g. word length and frequency) [159], which can lead to cross-language priming effects not due to semantics.

Previous fMRI studies based on univariate activation-based approaches did not show reliable differences in task-related (i.e. word generation, picture naming) hemodynamic activity across languages [67, 68, 160]. One limitation of these studies is that the experimental tasks and contrasts supposedly targeting semantic processing were often confounded by other untargeted orthographic/phonological processes [1]. Univariate fMRI-based priming studies [58, 161] have found some evidence for both language-shared and language-specific brain responses, but the role of strategic factors such as expectancy lists of prime-target relations could not be determined [158]. Strategies linked to expectancy lists (i.e. involving participants constructing a list of expected targets), alongside the use of long SOAs, may also alter the depth of processing [158]. Moreover, mass-univariate approaches are not best suited to identify whether or not semantic processing is mediated by a similar system across the different languages. The observation that a cortical area is activated in both languages does not imply that the brain representations are also similar. Two recent studies used multivariate pattern analysis (MVPA) to assess whether the brain activity patterns elicited by words in one language can predict the patterns of equivalent words in the other language [71, 72]. They found language-shared representations in well-known semantic substrates including the left parietal lobe, inferior frontal gyrus, and posterior temporal lobe.

A key limitation of the studies reviewed so far is that the factors underlying the generalization of semantic representations across languages remain to be determined. Critically, none of the previous MVPA studies noted above [71, 72] considered the depth of processing during the task. Here we operationalize the depth of processing based on the contrast between covertly reading a visual word (henceforth shallow processing) and mentally simulating the properties associated with the word concept (henceforth deep processing) based on the re-enactment of modality-specific representations. We note that while our manipulation of the depth of processing differs from the seminal experimental framework on 'levels of processing' [162] based on tasks targeting semantic vs. lower level phonemic/orthographic judgements, our experimental procedure is in keeping with different processing depths of processing; mental simulation is more likely to promote deeper semantic access, while the more shallow processing counterpart mainly taps onto phonological processing and rich semantic analyses is not mandatory.

Little research has examined the role of task-related factors on the brain representation of meaning. We here hypothesize that the depth of processing imposed by the task plays a critical role in the generalizability of semantic representations across languages. However, according to influential psycholinguistic models of word processing i.e. the Bilingual Integrated Activation model [56], the activation of language-shared representations may be independent of the depth of processing, and rather derived in parallel and non-selectively. Other theoretical accounts such as the perceptual symbols theory [22, 41] propose that semantic representations result from an implicit and automatic process of simulation in modality-specific sensory and action systems. This model therefore also predicts that semantic representations generalize across languages regardless of the depth of processing. Here we used fMRI-based MVPA to investigate how the depth of processing influences both within-language decoding and the generalization of semantic representations across languages in canonical substrates of the semantic network [1]. The cross-language generalization of the decoder was taken as a proxy for language-shared representations [156]. To pre-empt the results, we observed that while the decoding of the semantic category of words is significant within a given language regardless of the depth of processing, cross-language generalization of the brain representations of concepts was only found in the context of deeper levels of processing.

3.2 Materials and Methods

3.2.1 Participants

Thirty early and proficient Spanish-Basque bilinguals (mean age 24.2 ± 3.0 years; 19-34 years; 20 female) including twenty with Spanish as L1 were recruited through BCBL's own web portal specifically designed for this purpose: https://www.bcbl.eu/participa. They came from different educational backgrounds ranging from high school to postgraduate and professional training. All of them were healthy, had normal or corrected to normal vision, gave written informed consent prior to the experiment and were financially compensated with 20 euros for their time. The experiment was approved by the BCBL Ethics Review Board and conformed to the guidelines of the Helsinki Declaration.

All participants had acquired both languages before the age of 6. The age of acquisition of Spanish ($mean = 0.24 \pm 0.74$) was found to be statistically significantly lower ($t(29) = -2.60; p = 0.01$) than the age of acquisition of Basque (1.17 ± 1.61). Similarly, their reported performance in the two well known tests of language proficiency, i.e. LexTALE [118] - available for only 27 out of 30 participants - and BEST [119] - available for only 29 out of 30 participants - was also found to be statistically significantly higher (LexTALE: $t(26) = 5.46; p < 0.05$, BEST: $t(28) = 5.40; p < 0.05$) in Spanish (LexTALE: 94.54 ± 4.93, BEST: 99.36 ± 1.27) as compared to Basque (LexTALE: 86.56 ± 9.13, BEST: 89.76 ± 9.20). This shows that participants were more proficient in Spanish than in Basque.

Basque and Spanish are two very different languages with different roots. While Spanish is a romance language, Basque has unknown linguistic roots. It is an isolated pre-indo-european language. In addition, Basque holds many prominent linguistic differences with Spanish in the canonical word order in sentences regarding the subject, verb and object, morphology (Basque: agglutinative), syntax (Basque: ergative), and lexicon (many different vocabulary and non-cognates).

3.2.2 MRI Acquisition

A SIEMENS's Magnetom Prisma-fit scanner, with 3 Tesla magnet and 64-channel head coil, was used to collect, for each participant, one high-resolution T1-weighted structural image and ten functional acquisition runs each lasting for about 7 minutes. The proposed MR sequence was set up and run using SIEMENS's software Numaris/4 (version: syngo MR E11). In each fMRI run, a multiband gradient-echo echo-planar imaging sequence with acceleration factor of 6, resolution of $2.4 \times 2.4 \times 2.4$ mm^3, TR of 850 ms, TE of 35 ms, flip angle of 56 deg and bandwidth of 2582 Hz/Px was used to obtain 477 3D volumes of the whole brain (66 sagittal slices; FoV = 210 mm). The high resolution T1-weighted structural image covering the whole brain (resolution of $1.0 \times 1.0 \times 1.0$ mm^3, TR of 2530 ms, TE of 2.36 ms, flip angle of 7 deg) was collected after the fifth functional run using a fast 3D mprage sequence. The visual stimuli were projected on an MRI-compatible out-of-bore screen using a projector placed in the room adjacent to the MRI-room. A small mirror, mounted on the head coil, reflected the screen for presentation to the participants. The head coil was also equipped with a microphone that enabled the participants to communicate with the experimenters in between the runs.

3.2.3 Stimuli

A total of 16 words were used with 8 words per language. The Basque words were translational equivalents of Spanish words. Among 8 words, the 4 were living words including wolf, rooster, fox, and sheep, and the 4 were non-living words including candle, key, tube and mirror (for Spanish and Basque translations, see Figure 3.1). All the words were non-cognates and were balanced with respect to length and frequency (per million words; a standard measure independent of the corpus size) across categories (living and non-living) and languages ($t(7) = -1.16, p = 0.28$ for length and $t(7) = 0.28, p = 0.78$ for frequency per million; see Table 3.1 for details) based on the statistics provided by Espal (for Spanish; [120] and E-Hitz databases (for Basque; [121]). The requirement of length and frequency balancing across categories and languages put some constraints on the number of words; nevertheless the number finally selected was in keeping with previous studies of semantic decoding [71, 72, 122]. The semantic analysis of these words based on word embeddings i.e. word2vec (see A.7) show the non-living things to be more similar between them (light blue shade) as compared to the living things and there is a room for increased separation in the semantic space.

	Spanish		Basque	
	Living	Non-living	Living	Non-living
Length	4.5±0.58	4.75±0.96	4.5±0.58	5.25±0.96
Frequency	28.73±19.90	19.90±6.12	23.53±17.90	24.55±8.01

Table 3.1: The table shows mean word length and frequency per million of stimuli with respect to both languages and semantic categories. These statistics were gathered using Espal for Spanish and E-Hitz for Basque. It can be seen that they are balanced across categories and languages.

3.2.4 Experimental Procedure

Each trial began with a fixation period of 250 ms followed by a blank screen of 500 ms (see Figure 3.2). The target word, randomly drawn from a pool of 4 living and 4 non-living words (see 3.2.3), was presented for 1 s. Depending on a run's instructions (shallow or deep processing), the participants were supposed to either read and attend to the word, or to think about the characteristics of the living/non-living object it represented (e.g. its shape, its color etc.). Following a delay of 4 seconds, a red asterisk appeared at the center of the screen presented for a jittered time (see below) in which participants were instructed to do nothing. To ensure that the participants focused on the stimuli and the task, a maximum of two catch trials were set to appear at random points in each of the runs. These catch trials showed number words from among ZERO, ONE, and THREE in place of usual living/non-living words, and participants were supposed to respond by pressing any one of the four buttons on the fMRI response pad. The number TWO ("dos" in Spanish and "bi" in Basque) was not used due to different number of letters across languages. The total number of catch trials was kept equal across conditions.

To have as many trials as possible per each run, and at the same time maximize the separation between the brain activity corresponding to each of the trials, an event-related design was used and the time for which the asterisk stayed on the screen was jittered between 6 to 8s. This jitter was based on a pseudo-exponential distribution resulting in 50% of trials with the inter-trial interval of 6s, 25% with 6.5s, 12.5% with 7s and so on.

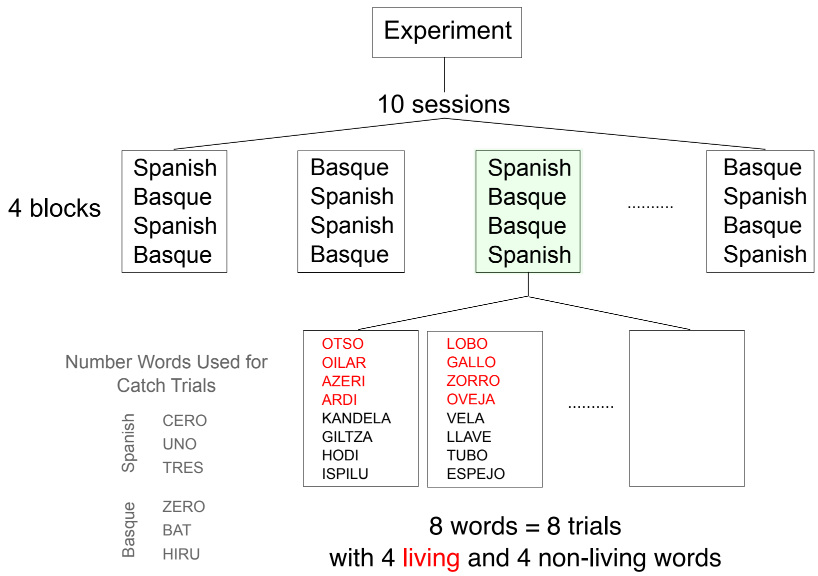


Figure 3.1: The figure summarizes the organization of runs, blocks and trials in the experiment. The experiment comprised of 10 runs with odd-numbered runs for shallow processing and even-numbered runs for deep processing. Each run was further subdivided into 4 language blocks (2 Spanish and 2 Basque). Each of these blocks was made up of 8 trials corresponding to single presentation of each of 4 living and 4 non-living words. The figure also shows the Spanish and Basque translations of both living/non-living words and number words.

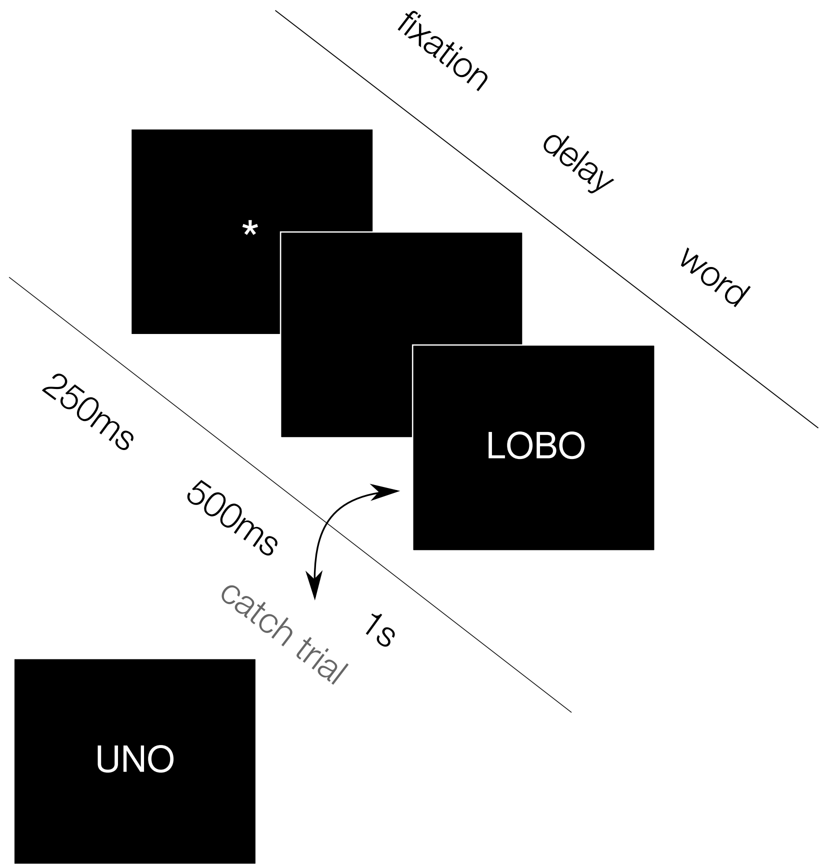


Figure 3.2: The figure illustrate the sequence of events on each trial. Following a fixation period, a word was presented for 1 s. The participants were supposed to either read and attend to the word or think about the living/non-living thing it represented. Next, following a delay of 4 s, a red asterisk appeared at the center of the screen and stayed there for a jittered interval of 6-8 s. To ensure that the participants were engaged, catch trials were placed at random points in each of the runs. These catch trials showed number words from among ZERO, ONE, and THREE in place of living/non-living words, and participants were supposed to respond by pressing a button.

Both instructions and stimuli were presented at the center of the screen, in white against black background and in all uppercase Arial font. The experiment was programmed using Psychopy [123] and is summarized in Figure 3.1. It comprised 10 runs (7 minutes each) and lasted for about 1.25 hours. In odd-numbered runs, participants were instructed to read and attend to the words (shallow processing), while in the even numbered ones, they were instructed to think about the characteristics of the living/non-living that the word represented (deep processing). Each fMRI run was subdivided into four language blocks with two Spanish (S) and two Basque (B) blocks, and the order of these blocks was counterbalanced across runs (SBSB, BSBS, and so on). In each of these blocks, eight words were presented (without repetition) in a random arrangement resulting in a total of thirty two trials per run.

3.2.5 MRI Data Preprocessing

The preprocessing of fMRI data was performed using FEAT (fMRI Expert Analysis Tool), a tool in FSL suite (FMRIB Software Library; v5.0). After converting all data from DICOM to NIfTI format using MRIConvert (`http://lcni.uoregon.edu/downloads/mriconvert`), the following steps were performed on each run's fMRI. To ensure steady state magnetisation, the first 9 volumes corresponding to the task instruction period were discarded; to remove non-brain tissue, brain extraction tool (BET) [124] was used; head-motion was accounted for using MCFLIRT [125]; minimal spatial smoothing was performed using a gaussian kernel with FWHM of 3mm. Next, ICA based automatic removal of motion artifacts (ICA-AROMA) was used to remove motion-induced signal variations [163] and this was followed by a high-pass filter with a cutoff of 60s. All the runs were aligned to a reference volume of the first run. All further analyses were performed in native BOLD space.

A set of 8 left-lateralized ROIs was pre-specified (see Figure 3.3) with 7 based on a meta-analysis of the semantic system by Binder et al. 2009 [1] and one anterior temporal lobe (ATL) due to its crucial role as a "semantic hub" [15, 37, 72]. So, the ROIs included: inferior parietal lobe (IPL), lateral temporal lobe (LTL), ventromedial temporal lobe (VTL), dorsomedial prefrontal cortex (dmPFC), inferior frontal gyrus (IFG), ventromedial prefrontal cortex (vmPFC), posterior cingulate gyrus (PCG) and anterior temporal lobe (ATL). First, automatic segmentation of the high-resolution structural image was obtained using FreeSurfer's automated algorithm `recon-all`. Next, `mri_binarize` was used to extract individual gray matter masks from aparc+aseg volume using corresponding label indices in FreeSurferColorLUT text file (`https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/AnatomicalROI`). And finally, after visually inspecting these in FSLView, they were transformed to each run's functional space using

FLIRT (7 DoF global rescale transformation). [125, 126] and were binarized (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT/FAQ).

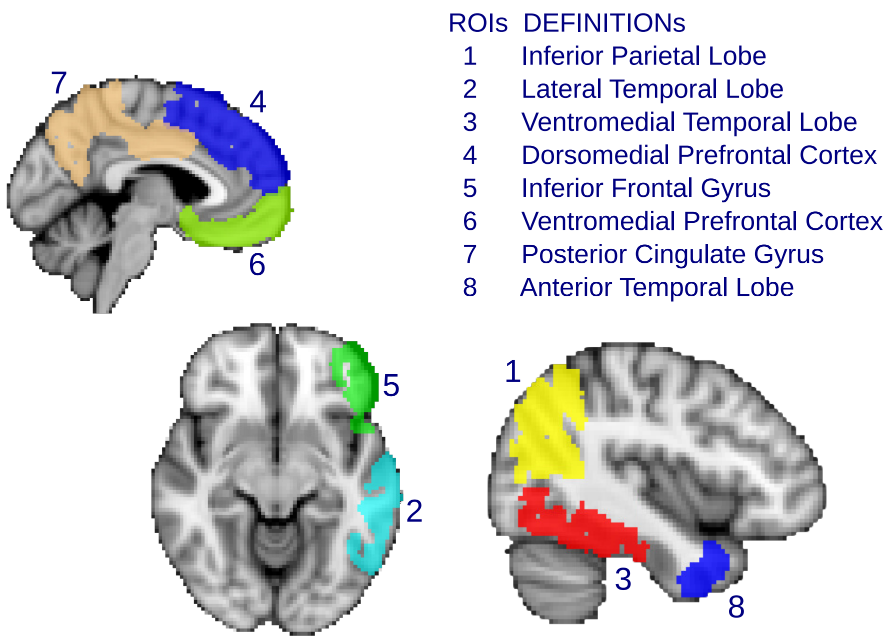


Figure 3.3: The figure shows the selected regions of interest projected on an MNI standard template image. These 8 left-lateralized areas were pre-specified and included inferior parietal lobe (IPL), lateral temporal lobe (LTL), ventromedial temporal lobe (VTL) including fusiform gyrus and parahippocampal gyrus, dorsomedial prefrontal cortex (dmPFC), inferior frontal gyrus (IFG), ventromedial prefrontal cortex (vmPFC), posterior cingulate gyrus (PCG) and anterior temporal lobe (ATL).

3.2.6 Multivariate Pattern Analysis

Multivariate pattern analysis was conducted using scikit-learn [127] and PyMVPA [128] libraries. Specifically, classification based on a supervised machine learning algorithm i.e. linear support vector machine [129], was used to evaluate whether multi-voxel patterns in each of the eight ROIs carry information related to the semantic category (living, non-living) of the word in each of the conditions. Within-language (or language-specific) decoding involved restricting the analysis to trials of a specific language (either Spanish or Basque) while cross-language (or language-independent) decoding entailed training the classifier on trials from one language and testing it on trials from another language. Both of these analyses were done separately for shallow and deep processing trials. Additional details related to the data preparation, feature selection, classification and statistics are presented in the following subsections.

3.2.6.1 Data Preparation

For each participant, the relevant time points or scans of the preprocessed fMRI data of each run were labeled with attributes such as word, category, language, and condition using Psychopy generated data files (CSVs). Invariant voxels (or features) were removed. These were the voxels/features whose value did not vary throughout the length of one run. If not removed, such features can cause numerical difficulties with procedures like z-scoring of features. Next, data from all ten runs were stacked and each voxel's time series was run-wise z-score normalized and linear detrended. Finally, following two recent cross-language generalization studies [71, 72], one example was created per trial by averaging the 4 volumes between the interval of 3.4 s and 6.8 s after the word onset, which corresponded to 1 second presentation of the word (see Figure 3.2). Importantly, this was the same in the shallow and deep processing conditions.

3.2.6.2 Pattern Classification

Linear support vector machine (SVM) classifier, with all parameters set to default values as provided by the scikit-learn package ($l2$ regularization, $C = 1.0$, $tolerance = 0.0001$), was used for both within- and cross-language decoding in both shallow and deep processing conditions. The following procedure was repeated for each ROI separately. To obtain an unbiased generalization estimate, following Varoquaux et al. 2016 [130] the data was randomly shuffled and resampled multiple times to create 300 sets of balanced train-test (80%-20%) splits. Since each example was represented by a single feature vector with each feature a mean of voxel intensities across the sub-interval of 3.4 s and 6.8 s (see 3.2.6.1), the length of a vector was equal to the number of voxels in the ROI. To further reduce the dimensionality of the data and thus reduce the chances of overfitting [131, 132], Principal Component Analysis (PCA) with all parameters set to default values as provided by the scikit-learn was used. Since the `n_components` argument was set to `None`, the number of components was chosen to be the smaller from among the number of samples (m) and features (n). In our case, the n was always greater than m, hence, the first m components were selected. The size of the data matrix after PCA was therefore $m \times m$. These components were linear combinations of the preprocessed voxel data and since none of the components was excluded, it was an information loss-less change of the coordinate system to a subspace spanned by the examples [133]. Features thus created were used to train the decoder, and its classification performance on the test set was recorded. This procedure was repeated separately for each of the 300 sets, and the mean of corresponding accuracies was collected for each of the participants. Note that PCA was performed on the training set; then the trained PCA was used to extract components in the test data and its classification performance was assessed. This procedure was repeated separately for each of the 300 sets, and the mean of corresponding accuracies was collected for each of the participants.

Our rationale to infer language-shared representations from the MVP classification analysis is based on the following logic: if a classier trained to discriminate stimulus classes in context A (or language A) generalises to discriminate the stimulus classes of previously unseen items in context B, there are grounds to argue that the underlying representations are similar across the two contexts and the level of similarity is proportional to the level of generalization performance of the classifier.

3.2.6.3 Statistics

To determine whether the observed decoding accuracy in a given ROI is statistically significantly different from the chance-level of 0.5 (or 50%), a two-tailed t-test was performed with p-values corresponding to each of the ROIs corrected for multiple comparisons using a false discovery rate (FDR) method. To get the empirical estimate of chance-level, we ran the classification tests while randomly permuting over the category labels. The chance-level was computed across participants, ROIs, classification problems (within and cross-language) and conditions. For each case, 300 permutations were performed and the mean and standard deviation of the collected permutation scores was calculated across participants. For all ROIs, and classification problems, the chance-level was consistently found to be centered around 0.5. All effect sizes are reported as *mean effect size*$\pm$*standard* *error*, *t(degrees of freedom)=t-value, p-value* across all participants.

3.3 Results

3.3.1 Behavioral

To ensure that participants were attending to the items during the task, a few catch trials were randomly presented at different points in each run. These trials showed number words and required a response via button press. Further details related to the participants and procedure are provided in 3.2. To ensure equal treatment of both conditions, the total number of catch trials ($mean = 6.8 \pm 1.6$) was kept equal in both shallow and deep processing runs. Catch trial data from two initial participants could not be obtained due to a technical issue. The proportion of correct responses on catch trials was 0.90 ± 0.13 in the shallow processing, and 0.93 ± 0.12 in deep processing conditions, which did not differ ($t(27) = 0.87, p = 0.39$), hence showing that participants were equally engaged with the task in both conditions.

3.3.2 FMRI-based MVPA Results

For each participant, we performed MVPA in 8 well-known left-lateralized semantic ROIs (see Figure 3.3). We asked whether shallow processing is sufficient for decoding the word semantic category within a given language and also to activate semantic representations that generalize across languages; or, whether higher depth of processing is needed for such cross-language generalization. Specifically, linear support vector machine (SVM) was used for classification of the semantic category in all ROIs in both shallow and deep processing conditions. Two different classification analyses were performed, namely within-language decoding and cross-language generalization. Both of these were performed separately for each of the conditions on each subject, and were restricted to eight pre-specified ROIs based on a prior meta-analysis [1]. To determine whether the observed decoding accuracy in a specific ROI and condition is statistically significantly above chance, a two-tailed t-test was performed. All t-tests reported below were corrected for multiple comparisons using FDR method.

3.3.2.1 Within-language Decoding

Within-language decoding was restricted to one language at a time whereby 80% of trials of that language were used to train the SVM-based classifier and the remaining 20% to test the learned model. Figures 3.4 and 3.5 therefore present the summary statistics of the ROIs for both shallow and deep processing conditions within Spanish and Basque respectively. It can be seen that in both the shallow and deep processing conditions, the decoding of the semantic category (living/non-living) was found to be statistically significantly above chance in almost all pre-specified ROIs (see Figures 3.4, 3.5 and Supplemental Results A.1 for statistics).

Deep processing also resulted in relatively higher decoding performance relative to the shallow processing condition in some of the ROIs. Specifically, deep processing was found to improve within-language decoding in IPL ($p = 0.004$), VTL ($p = 0.02$), and PCG ($p = 0.002$) for Spanish and IPL ($p = 0.001$), VTL ($p = 0.008$) and IFG ($p = 0.04$) for Basque. It can also be seen that an exception to this was ATL where decoding in the shallow condition was found to be higher than that in deep condition ($p = 0.002$ for Spanish, $p = 0.75$ for Basque).

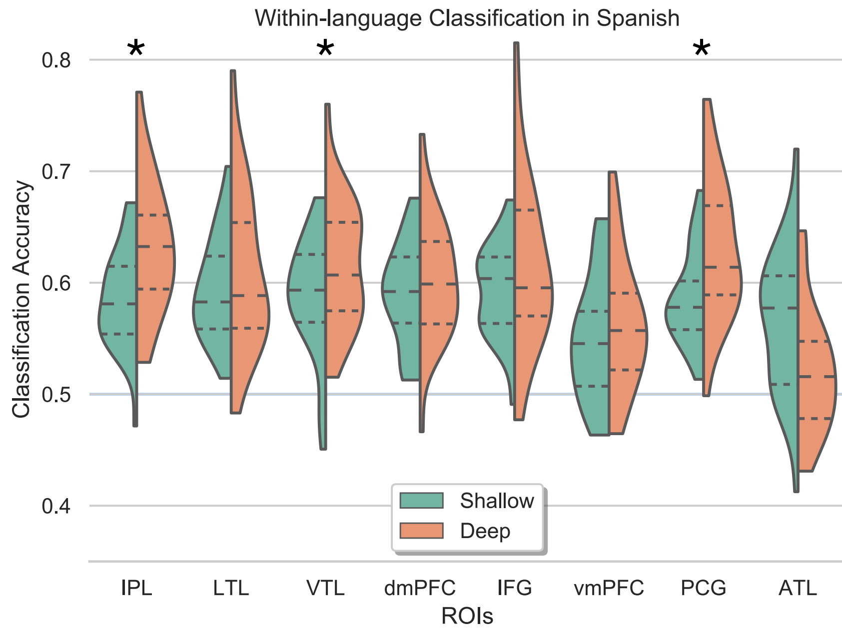


Figure 3.4: The figure shows summary statistics of the ROIs for within-language decoding in Spanish. It can be seen that the decoding was above chance in all ROIs in both conditions. The three dotted lines inside each violin are the quartiles. The black asterisks mark ROIs that showed statistically significant improvement in decoding accuracy in deep as compared to shallow processing condition. The p-values were corrected for multiple comparisons.

We conducted further control analyses to address the following points. First, it may be argued that the decoding accuracy in the within-language classification could reflect low-level features of the items given that the same words (though different examples) were used in training and testing the classifier. We believe this is an unlikely explanation because we controlled for linguistic properties (i.e. length and frequency) of the items. Further, classification accuracy was quite distributed across the ROIs, including high-level semantic ROIs. Nevertheless, the within-language decoding analyses were re-run with the classifier trained on all words but one and tested on the left-out word. Similar results were observed, although the level of decoding accuracy was somehow weaker across ROIs and the within-language decoding was most evident in Spanish relative to Basque (see Supplemental Results A.3). It is possible that any seemingly stronger effect in Spanish may be due to the fact that most of our participants had Spanish as the first language. However, since this is not the focus of the study, this issue will not be discussed further. In summary, these results show that within-language decoding did not reflect low-level features of the words. Note that this issue does not apply in the case of cross-language generalization, which is the focus of the present study.

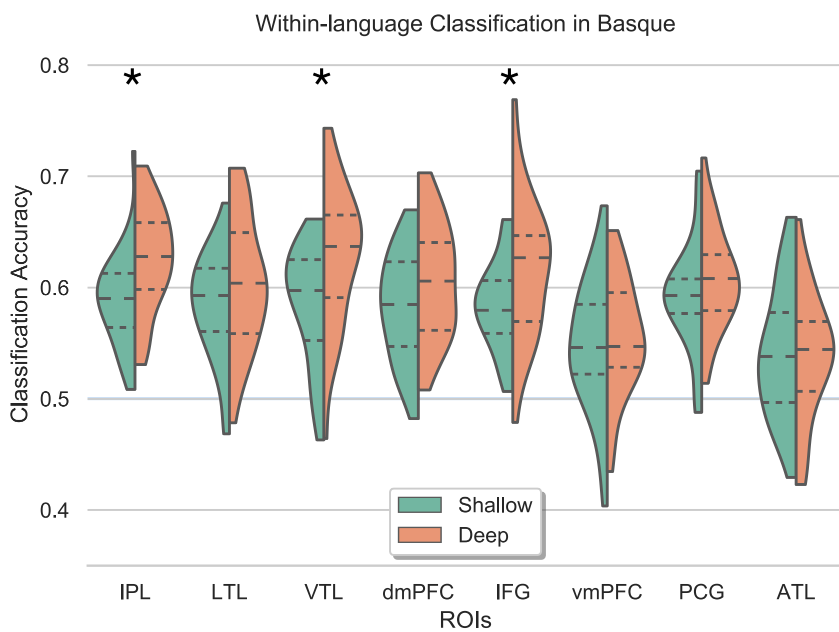


Figure 3.5: The figure shows summary statistics of the ROIs for within-language decoding in Basque. It can be seen that the decoding was above chance in all ROIs in both conditions. The three dotted lines inside each violin are the quartiles. The black asterisks mark ROIs that showed statistically significant improvement in deep as compared to shallow processing condition. The p-values were corrected for multiple comparisons.

3.3.2.2 Cross-language Decoding

Cross-language generalization involved training the decoder on the examples of one language (training language) and testing it on the examples of the other language (test language). Figures 3.7 and 3.6 present summary statistics of the ROIs for Spanish to Basque and Basque to Spanish generalization respectively in both shallow and deep processing conditions. It can be seen that in the shallow processing condition, the cross-language generalization from both Spanish to Basque and Basque to Spanish was not different from chance-level in all pre-specified ROIs (see Supplemental Results A.2 for additional details).

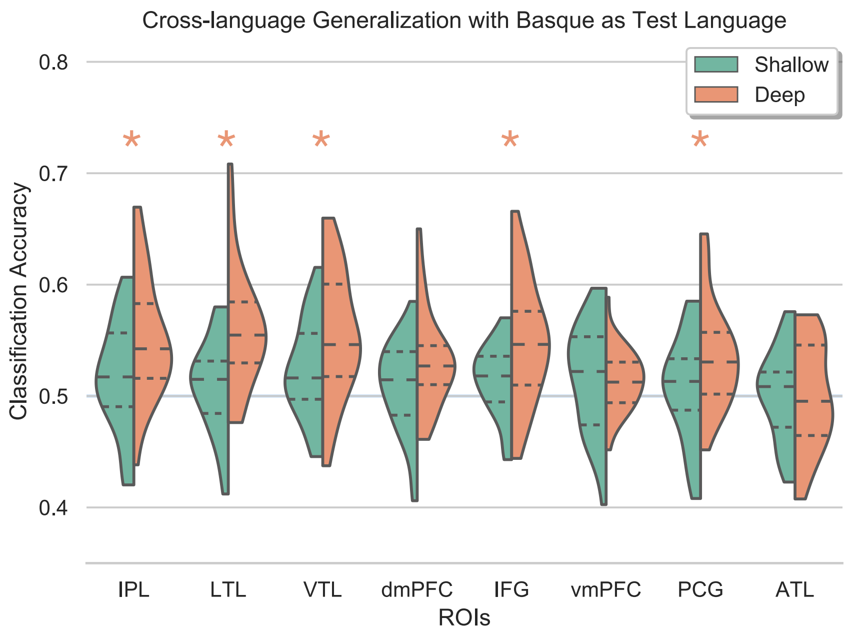


Figure 3.6: The figure shows summary statistics of the ROIs for cross-language generalization from Spanish to Basque in both shallow and deep processing conditions. It can be seen that whereas the generalization was not different from chance in all ROIs in the shallow condition, it was statistically significantly above-chance and better than shallow condition in deep condition in five out of eight ROIs including IPL, LTL, VTL, IFG and PCG. The three dotted lines inside each violin are the quartiles. The orange asterisks mark ROIs where cross-language generalization in deep was found to be statistically significantly above chance and better than shallow condition. The p-values were corrected for multiple comparisons.

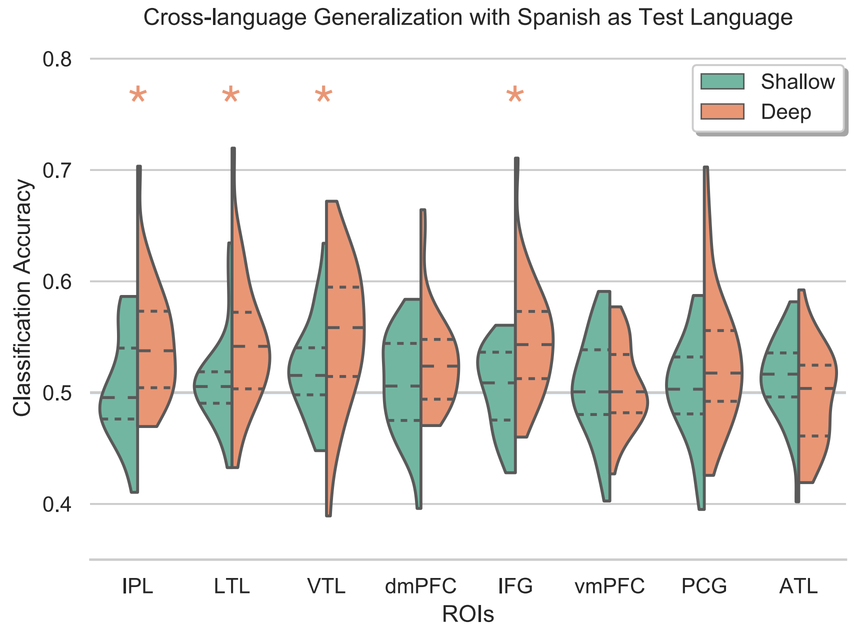


Figure 3.7: The figure shows summary statistics of the ROIs for cross-language generalization from Basque to Spanish in both shallow and deep processing conditions. It can be seen that while the generalization was not different from chance in all ROIs in the shallow condition, it was statistically significantly above-chance and better than shallow condition in deep condition in four out of eight ROIs including IPL, LTL, VTL and IFG. The three dotted lines inside each violin are the quartiles. The orange asterisks mark ROIs where cross-language generalization in deep was found to be statistically significantly above chance and better than shallow condition. The p-values were corrected for multiple comparisons.

In the deep processing condition on the other hand, the Spanish to Basque generalization was found to be statistically significantly above-chance and better than shallow condition (FDR corrected for multiple comparisons) in five out of eight ROIs including: IPL ($55.18 \pm 5.27, t(29) = 5.29, p = 2.99e-05$), LTL ($55.84 \pm 5.35, t(29) = 5.88, p = 1.78e-05$), VTL ($55.45 \pm 5.49, t(29) = 5.34, p = 2.99e-05$), dmPFC ($53.12 \pm 4.25, t(29) = 3.95, p = 0.0006$), IFG ($54.89 \pm 5.33, t(29) = 4.94, p = 6.00e-05$), vmPFC ($51.47 \pm 2.74, t(29) = 2.89, p = 0.008$), PCG ($53.57 \pm 4.62, t(29) = 4.15, p = 0.0004$), ATL ($50.01 \pm 4.69, t(29) = 0.02, p = 0.99$). Similarly, Basque to Spanish generalization was found to be statistically significantly above chance and better compared to shallow condition (FDR corrected for multiple comparisons) in four out of eight ROIs including: IPL ($54.50 \pm 5.12, t(29) = 4.74, p = 0.0002$), LTL ($54.47 \pm 5.72, t(29) = 4.21, p = 0.0005$), VTL ($55.34 \pm 6.36, t(29) = 4.52, p = 0.0003$), dmPFC ($53.05 \pm 4.38, t(29) = 3.75, p = 0.001$), IFG ($54.78 \pm 5.06, t(29) = 5.08, p = 0.0002$), vmPFC ($50.55 \pm 3.65, t(29) = 0.82, p = 0.48$), PCG ($53.03 \pm 5.95, t(29) = 2.74, p = 0.01$), ATL ($49.64 \pm 4.28, t(29) = -0.45, p = 0.65$)). Notably, above-chance cross-language generalization in the deep condition was not restricted to ROIs that showed superior within-language decoding as compared to the shallow condition (see Figures 3.4 and 3.5). We come back to this point in the Discussion.

The above results clearly show that cross-language generalization was stronger in the deep compared to the shallow processing condition. Because parametric statistical tests were used, additionally Shapiro-Wilk tests were run to check the normality assumption in the data. The results showed that normality assumption held in our dataset. Additionally, we also ran non-parametric statistical tests i.e. Wilcoxon signed-rank tests and found a similar pattern of results to those obtained using parametric t-tests. Furthermore, we also ran Bayesian analyses with all parameters set to default values in the JASP statistical package [164, 165]) to assess the extent of the evidence for the null hypothesis in the cross-language generalization in the shallow condition (see Supplemental Table A.5). The results here showed that evidence for the null hypothesis in the shallow condition ranged from moderate to anecdotal in all of the ROIs. While this could be interpreted as non-conclusive evidence for the absence of generalization in the shallow processing case, the key observation is that generalization is far stronger in the deep relative to the shallow processing condition. The evidence for the alternative hypothesis in the deep processing context was found to be extreme in most of the ROIs (see Supplemental Table A.6).

Given the evidence for ATL involvement as a semantic hub, we performed some additional analysis in the ATL. The results presented above showed that while significant decoding in the ATL was found in the shallow context, there was no evidence of cross-language generalization even during deep processing. This result was obtained with a mask of the ATL based on Freesurfer anatomical segmentation and is in keeping with the study of Damasio et al. 1996 [37] and Correia et al. 2014 [72] in which cross-language generalization was found. However, there is a further, relatively more posterior ATL area that was also implicated as a multi-modal semantic hub (see Chen et al. 2017 [2]). To re-run the decoding analysis on this area, we derived a 6 mm mask for each subject in native space based on registration from the corresponding MNI coordinates (-39, 18, -30), which lie between ROIs 3 and 8 in Figure 3.3. We found above-chance within-language decoding in both shallow (Spanish: $52.78 \pm 5.82, t(29) = 2.57, p = 0.02$; Basque: $53.48 \pm 6.00, t(29) = 3.12, p = 0.004$) and deep (Spanish: $54.60 \pm 4.48, t(29) = 5.53, p = 7.00e{-}06$; Basque: $54.14 \pm 5.08, t(29) = 4.40, p = 0.0001$) conditions with no significant differences between them ($p = 0.20$ for Spanish, and 0.63 for Basque), chance-level cross-language generalization was found in both conditions (shallow: $51.19 \pm 3.82, t(29) = 1.67, p = 0.19$ for Spanish to Basque and $50.66 \pm 3.52, t(29) = 1.01, p = 0.56$ for Basque to Spanish generalization; deep: $50.16 \pm 4.15, t(29) = 0.20, p = 0.95$ and $50.03 \pm 4.21, t(29) = 0.03, p = 0.97$).

Next, we explore several factors that may account for the apparent absence of cross-language generalization in the shallow condition.

We wondered whether cross-language generalization in the shallow condition may be related to inter-individual differences in language proficiency scores in BEST and LeXTALE tests. Hence, we assessed the correlation of language proficiency and cross-language decoding accuracy in the different ROIs. Specifically, we expected that balanced bilinguals, namely, participants with minimal difference in Spanish and Basque proficiency scores would display increased cross-language generalization accuracy (mean generalization scores across Spanish to Basque and vice versa). However, we did not find reliable evidence in support of this hypothesis (see Supplemental Results A.6).

It could be argued that the sub-interval of 3.4s-6.8s may not be the most optimal choice for creating examples (see 3.2.6.1). As mentioned above, the choice of this time window was based on previous cross-language generalization studies [71, 72] and standard guidelines in the field of fMRI-based multivariate pattern decoding [131]. However, we also re-ran the whole analysis taking the average of 2 volumes across the sub-interval of 4.25 s and 5.95 s. We found both within-language and cross-language generalization to be similar to those obtained using the sub-interval of 3.4 s and 6.8 s. Specifically, cross-language generalization was again found to be at chance-level in all ROIs in the shallow condition.

It could also be argued that information critical for cross-language generalization in the shallow condition is stored in spatially distributed, remote brain areas [166]. Given that our ROI-based approach restricted the MVP analysis to one ROI at a time, it remains possible that significant cross-language generalization in the shallow condition is observed with a bigger ROI. To investigate this, we combined the data from all eight ROIs and repeated the analysis in the shallow condition. We found above-chance within-language decoding (Spanish: $59.12 \pm 4.47, t(29) = 10.98, p = 7.61e{-12}$; Basque: $58.24 \pm 4.88, t(29) = 9.10, p = 5.36e{-10}$), but cross-language decoding was not different from chance during Spanish to Basque generalization ($51.37 \pm 4.77, t(29) = 1.54, p = 0.13$) and Basque to Spanish generalization ($51.19 \pm 4.46, t(29) = 1.43, p = 0.16$).

Conversely, it could also be argued that the pre-specified ROIs were relatively large and the PCA merged features that were irrelevant for further classification analysis [167]. This could be suggested as one possible reason for chance-level cross-language generalization in the shallow condition. In an attempt to address this point, the 8 ROIs were further subdivided into 15 more fine-grained ROIs based on individual anatomically segmented masks from Freesurfer (i.e. including inferior parietal lobe, inferior temporal lobe, medial temporal lobe, fusiform gyrus, parahippocampal gyrus, superior frontal gyrus, pars opercularis, pars orbitalis, pars triangularis, lateral orbitofrontal cortex, medial orbitofrontal cortex, posterior cingulate gyrus, precuneus, and anterior temporal lobe). Then, the same MVP analysis was repeated. However, cross-language generalization in the shallow condition was not different from chance in all ROIs for both Spanish to Basque and Basque to Spanish generalization, while crucially generalization was significantly above chance in the deep condition in a number of ROIs located in the anatomical spaces of the 8 ROIs described above (see Supplemental Results A.5).

3.4 Discussion

An important question in psychology and neuroscience is whether the acquisition of different languages is integrated within the same neurocognitive system and include shared/overlapping representations, or whether different languages are represented in separate brain systems. Previous investigations did not address the factors that underlie the generalization of semantic representations across languages. Hence it remained to be determined whether and how semantic representations generalise across languages. This fMRI study provides novel insights into this issue by uncovering how the depth of processing during semantic tasks influences within-language decoding of word category and cross-language generalization based on multivoxel patterns of BOLD responses in putative substrates of the semantic network.

We found that the semantic category of words could be significantly decoded above chance levels when both Spanish and Basque languages were considered separately in all pre-specified semantic areas based on a prior meta-analysis [1]. This happened even under shallow processing conditions when participants were merely asked to attend and read the words. However, the decoding performance was significantly better in deep compared to shallow processing in IPL, VTL, and PCG for Spanish and IPL, VTL, and IFG for Basque. The superior decoding performance in the deep relative to shallow processing condition aligns with other recent observations in our laboratory [168] and indicates that the task requirement had an impact on the brain representation of meaning.

Cross-language generalization was not different from chance in all ROIs during shallow processing conditions (see also 2). Only in the context of deep information processing did brain activity patterns reliably generalize from Spanish to Basque in several brain regions (from Spanish to Basque: IPL, LTL, VTL, dmPFC, IFG, and PCG; from Basque to Spanish: IPL, LTL, VTL and IFG) known to be involved in semantic processing. For instance, the left IPL has been found to allow cross-language generalization in fMRI studies using visual [71], auditory word comprehension with concrete nouns [72] and also during narrative comprehension task [156]. PCG, and dmPFC have previously been found in cross-language generalization with visual stimuli [71, 156] but not in those using auditory stimuli [72]. Similarly, LTL and VTL have been found to carry patterns that generalize across languages in studies using visual word comprehension [71] as well as production tasks [169, 170].

It is worth noting that cross-language generalization in the deep condition was also found in ROIs which showed no difference in within-language decoding as function of the depth of processing. Specifically, multivoxel patterns in the lateral temporal lobe and dorsomedial prefrontal regions contained information that generalized across languages only in the context of a higher depth of processing but not during shallow processing, despite within-language decoding accuracy was the same in deep and shallow contexts. This pattern of results indicates that cross-language generalization in the deep processing case is not merely due to the increased signal to noise ratio of the multivoxel patterns corresponding to living and non-living items or merely based on modality-specific representations triggered by mental (e.g. visual) imagery processes occurring more strongly during deep relative to shallow processing [168]. Our results also indicate that language-independent neural representations of semantic knowledge may not be easily generated during bottom-up information processing (i.e. automatically) but may require top-down strategic control processes [171] such as those triggered during deep information processing and mental simulation.

The influential hub-and-spoke model suggests that sensory-motor representations of a concept are encoded in modality-specific brain regions (spokes), yet, unified and amodal representations are formed within a single transmodal hub in anterior temporal lobes (ATL). On the other hand, the distributed-only model suggests that the higher-order generalizations from modality-specific (or language-specific) to amodal (or language-independent) semantic representations is not confined to a single semantic hub, rather distributed multiple brain regions are involved [15, 172]. In our study, we found significant within-language decoding in ATL, yet cross-language generalization was not observed in this region in both shallow and deep conditions. These null results however must be taken with caution given that ATL is well-known to have susceptibility-induced signal dropout issues, and also considering the amount of evidence in the favour of the key role of ATL as a multi-modal semantic hub [173]. The critical finding however is that the cross-language generalization was found in multiple substrates of the semantic network. This is in keeping with previous neuroimaging studies [71, 72], though here we revealed the critical role of the depth of processing. We propose that the depth of information processing triggered the global sharing of information across a distributed set of brain areas implicated in semantic representation and this supported cross-language generalization. We suggest that the present results are in keeping with distributed-only views of semantic processing [15].

We observed significant decoding of semantic category in inferior parietal, medial and inferior temporal and inferior frontal regions. These cortical association areas, also known as a "transmodal cortex" [174], are thought to play a critical role in higher-order semantic processing [173]. Although the specific role of inferior frontal cortex still remains a topic of debate, previous studies indicate that it is not involved in the storage of semantic knowledge as such, but in semantic control [175, 176, 177]. In our study, word meaning could be decoded from patterns of activity in inferior frontal gyrus, namely, pars opercularis and pars triangularis, both within-, and also cross-languages. These results implicate this region in semantic representation (see also [71, 122, 168]). It is typically assumed that bilinguals are constantly switching between the two languages, selecting one and inhibiting the other based on task goals. However, it is hard to explain the within- and cross-language decoding of semantic categories based on this language switching account and semantic control view.

The present results have ramifications for psycholinguistic models of visual word recognition e.g. BIA+ [56]. These models implement word processing in a purely bottom-up manner with parallel and non-selective (i.e. language independent) activation of linguistic codes not just at the level of semantics but orthography and phonology too. We propose that such models need to be revised to incorporate the influence of top-down factors related to the depth of processing. Our results indicate that non-selective access to word meaning across languages is not mandatory or intrinsic property of the semantic system. Instead, our results are in keeping with the view that depending on the depth of processing, the extent of parallel and non-selective access can be modulated. For instance, studies that did not encourage high depth of processing only found evidence for selective access [178, 179]. More research is however needed to elucidate the extent to which the depth of processing shapes how bilinguals access semantic representations, namely, the extent to which different language representations for a given concept are co-activated in parallel [180] or whether, according to BIA+, bilinguals access to the lexical and semantic representation is delayed in the second language compared to the first language [181]. Furthermore, here we only used eight words per language in order to match them as much as possible in linguistic factors, however, the limited number of words imposes constraints on the scope of inferences that can be drawn about the neurocognitive architecture of the semantic system across different languages. Future studies using a larger corpus of words, time-resolved electrophysiology and computational models are needed to pinpoint the effect of the depth of processing and other task-related factors on the brain dynamics for accessing semantic representations in different languages. Ongoing work in the lab is being directed to test this view.

An additional limitation of the present study may relate to the high sampling rate used (multiband acceleration factor of 6), which might have led to signal loss in some regions and geometric distortion that can affect the anatomical registration of the functional images. No field maps were obtained to correct for potential field inhomogeneities. However, inspection of our images did not reveal greater distortions compared to standard (i.e. no multiband) acquisitions. Additional research is needed to achieve a comprehensive evaluation of the relationship between acquisition parameters (MB factors, in-plane acceleration, voxel size, TR, flip angle) and MVPA decoding results, and benefits in event-related designs with short trial event have already been demonstrated through a comparison of multiband 2 and 3 (see [182], also [183]). Of note, however, the level of decoding performance in the present study was similar to previous MVPA decoding studies that used similar paradigms with standard MRI sequences [71, 72, 122].

Chapter 4

The Role of Lexico-Semantic Factors in the Cross-Language Generalization of Semantic Representation

4.1 Introduction

Neuroimaging studies of bilinguals show that the brain activity patterns created during semantic processing in the first language generalize to those created during semantic processing in the second language ([71, 72]). Specifically, these functional magnetic resonance imaging (fMRI) based multivariate pattern analysis (MVPA) studies show certain areas of the brain (including left parietal lobe, inferior frontal gyrus and posterior temporal lobe) to have above-chance cross-language generalization, suggesting the presence of language-shared or language-independent semantic representations. A key limitation of these studies however is that they assume the activation of language-shared semantic representations to be automatic and therefore do not consider potential factors that may mediate the generalization of these representations across languages.

The Chapter 2 of this thesis presented an fMRI-based MVPA study, involving Spanish-Basque bilinguals, which used masked animal and non-animal words as stimuli, and investigated if the level of conscious awareness influences the cross-language generalization of semantic representations from Spanish to Basque and vice versa. Masking, coupled with a relatively low luminance, resulted in stimuli with a very low signal to noise ratio, leading to very low within-language decoding and consequently only chance-level cross-language generalization due to floor effect. It is with a separate analysis involving only the participants with relatively high within-language decoding that the level of conscious awareness was found to weakly influence the cross-language generalization of semantic representations from Spanish to Basque. Similarly, the Chapter 3 presented another fMRI study, also involving Spanish-Basque bilinguals, with the same stimuli. This study investigated if such task-related factors as the depth of processing affects the cross-language generalization from Spanish to Basque and vice versa. This study consisted of two conditions i.e. shallow processing: participants were instructed to read the animal/non-animal words and deep processing: they were instructed to read the words and think about the concepts that they represented (similar to [71]).

Both of these studies used a small set of concrete and familiar words as stimuli (i.e. 8 words per language, 4 per each category). This constrains the type of general inferences that can be drawn about the neurocognitive architecture of the semantic system across different languages. The proposed fMRI study fills this gap by employing a much larger and diverse set of words with varying levels of frequency (highly frequent words are usually highly familiar too) and concreteness. Since no prior study has looked into the effect of such lexico-semantic factors as frequency and concreteness, the proposed study used fMRI-based MVPA to investigate how these factors influence the cross-language generalization of meaning in canonical substrates of the semantic network [1].

When it comes to the semantic processing of abstract words, an important question is *how semantic processing of abstract words is different from the semantic processing of concrete words?*. Previous behavioral and neuroimaging studies show the underlying mechanisms to be doubly dissociated i.e. the neural representations of concrete and abstract concepts are at least partially distinct. Specifically, in lesion studies, patients with brain damage have been seen to make more errors for abstract than concrete items [184, 185]. On the other hand, behavioral studies, in general, demonstrated faster and more accurate processing of concrete than abstract words [186, 187]. Specifically, the cross-language translation priming effects were found to be stronger from L1 to L2 than from L2 to L1 for concrete than abstract word pairs [52]. It was argued that this maybe due to difference in imageability with concrete concepts having higher imageability than abstract concepts [188]. Similarly, the cross-language association priming effects were also found to be stronger for concrete than abstract words [47]. The authors argued that the cross-language semantic overlap is dependent on the number of shared features and is lower for abstract concepts as they are often used in more different contexts across languages. Note however that Francis and Goldmann (2011) [55] reported contrasting findings i.e. similar and symmetric cross-language priming effects for abstract and concrete words, suggesting a complete overlap in semantic representations across languages, independent of the level of concreteness. It is however impossible to rule out methodological issues associated with these cross-language priming studies [157, 158].

Different theories of meaning representation have considered the role of mental imagery as an important factor. On the one hand, perceptual symbol systems theory hypothesizes that all concepts, independent of the level of concreteness, are grounded in perception and action [4, 7, 22]. On the other hand, the dual-coding theory postulates a common verbal representation for both abstract and concrete concepts with an additional mental imagery process for concrete concepts [188].

Word frequency is considered as a strong predictor of processing efficiency [189]. High frequency words are usually more familiar to people and are thus processed faster than low frequency words. This is as true for the low-level lexical as for high-level semantic processing. The key relevant question is the effect of frequency and concreteness in the neural representation of concepts and how they generalize across languages.

Accordingly, the proposed fMRI-based MVPA study presented Spanish-Basque bilinguals with Spanish and Basque stimuli (translational equivalents), comprising of animal/non-animal words, with varying levels of frequency and concreteness. Cross-language generalization of semantic category was then used to localize the cross-language overlap of semantic representations.

4.2 Materials and Methods

4.2.1 Participants

Twenty early and proficient Spanish-Basque bilinguals (mean age 24.7±3.5 years; 20-31 years; 13 female) including nine with Spanish as L1 were recruited through BCBL's own web portal specifically designed for this purpose: `https://www.bcbl.eu/participa`. They came from different educational backgrounds ranging from high school (5/20) to postgraduate (1/20) and professional training (3/20). All of them were healthy, had normal or corrected to normal vision, gave written informed consent prior to the experiment and were financially compensated with 40 euros (20 euros/day) for their time. One of the participants was excluded before fMRI-based MVPA analysis due to excessive motion. The experiment was approved by the BCBL Ethics Review Board and conformed to the guidelines of the Helsinki Declaration.

All participants had acquired both languages before the age of 10. The age of acquisition of Spanish ($mean = 1.57 \pm 2.85$) was not found to be statistically significantly different ($p = 0.27$) from the age of acquisition of Basque (0.68 ± 1.20). However, as far as their reported performance in the two well known tests of language proficiency, i.e. LexTALE [118] and BEST [119] is concerned, while the

LexTALE score was not found to be statistically significantly different ($p = 0.50$) between Spanish (LexTALE: 95 ± 0) and Basque (LexTALE: 89.79 ± 9.06), the BEST score was found to be statistically significantly higher ($p = 0.003$) in Spanish (99.43 ± 1.05) as compared to Basque (93.20 ± 7.37). This shows that participants were relatively more proficient in Spanish than in Basque.

Basque and Spanish are two very different languages with different roots. While Spanish is a romance language, Basque has unknown linguistic roots. It is an isolated pre-indo-european language. In addition, Basque holds many prominent linguistic differences with Spanish in the canonical word order in sentences regarding the subject, verb and object, morphology (Basque: agglutinative), syntax (Basque: ergative), and lexicon (many different vocabulary and non-cognates).

4.2.2 MRI Acquisition

A SIEMENS's Magnetom Prisma-fit scanner, with 3 Tesla magnet and 64-channel head coil, was used to collect, for each participant, one high-resolution T1-weighted structural image and sixteen functional acquisition runs (eight per day) each lasting for about 7 minutes. The proposed MR sequence was set up and run using SIEMENS's software Numaris/4 (version: syngo MR E11). In each fMRI run, a multiband gradient-echo echo-planar imaging sequence with acceleration factor of 6, resolution of $2.4 \times 2.4 \times 2.4$ mm^3, TR of 850 ms, TE of 35 ms, flip angle of 56 deg and bandwidth of 2582 Hz/Px was used to obtain 477 3D volumes of the whole brain (66 sagittal slices; FoV = 210 mm). The high resolution T1-weighted structural image covering the whole brain (resolution of $1.0 \times 1.0 \times 1.0$ mm^3, TR of 2530 ms, TE of 2.36 ms, flip angle of 7 deg) was collected after the fifth functional run using a fast 3D mprage sequence. The visual stimuli were projected on an MRI-compatible out-of-bore screen using a projector placed in the room adjacent to the MRI-room. A small mirror, mounted on the head coil, reflected the screen for presentation to the participants. The head coil was also equipped with a microphone that enabled the participants to communicate with the experimenters in between the runs.

4.2.3 Stimuli

We selected 128 nouns with non-cognate translational equivalents across Spanish and Basque for a total of 256 words (see Appendix B.1). Half of the nouns were animal and half non-animal. Within each of these semantic classifications, we orthogonally manipulated lexical frequency and concreteness to create four conditions based on a median split on each variable (see Table 4.1 and Figure 4.3). This produced 16 total categories with 16 stimuli per category. Lexical frequency (log frequency per million) values were extracted from the B-Pal and E-Hitz corpora for Spanish and Basque, respectively [121, 190]. Concreteness values were obtained from the BaSp translation database [191]. Frequency, concreteness, length and orthographic neighborhood (based on Coltheart's N; [192]) were matched across languages, semantic classifications, and median splits on frequency and concreteness ($p > 0.14$; see Table 4.2).

Frequency	Concreteness	Frequency	Concreteness
High	High	1.6±0.3	5.0±0.5
	Low	1.7±0.6	3.6±0.5
Low	High	1.0±0.2	5.1±0.4
	Low	0.9±0.2	3.5±0.4

Table 4.1: The table shows means and standard deviations of frequency (log per million) and concreteness (scale of 1-7) across median splits of each variable.

	Spanish		Basque	
	Animal	Non-animal	Animal	Non-animal
Length	7.2±1.4	6.8±1.5	7.2±1.8	6.9±1.6
Frequency	1.3±0.4	1.3±0.4	1.3±0.5	1.3±0.6
Concreteness	4.4±0.8	4.3±0.9	4.5±0.8	4.3±0.8
Neighborhood	1.1±1.2	1.2±1.7	1.3±2.3	1.3±2.2

Table 4.2: The table shows mean frequency (log per million), concreteness (scale of 1-7), length (number of characters), and orthographic neighborhood (Coltheart's N) with corresponding standard deviations with respect to both languages and semantic categories. These statistics were gathered using B-Pal for Spanish and E-Hitz for Basque.

4.2.4 Experimental Procedure

Each trial began with a fixation period of 250 ms followed by a blank screen of another 500 ms (see Figure 4.1). The target word, randomly chosen from a pool of 256 words (128 per language, 32 per condition; see Figure 4.3), was presented for 1s and was followed by a response period of 1.5 s. During this period, the participants were asked one question i.e. which semantic category does the word belong to, animal (A) or non-animal (nA)? To eliminate the effect of motor response difference on the choice of a semantic category, the mapping between choice and response button was randomly assigned on each trial. So, for some trials, A was on the right with nA on the left of the response screen, while for others, A was on the left with nA on the right. Participants were instructed to make their choice between left (i.e. button 1) and right (i.e. button 2) buttons based on the text displayed ("A nA"/"nA A") during the response period. This response text was finally replaced by an asterisk which stayed on the screen for a jittered interval of 6-8 s.

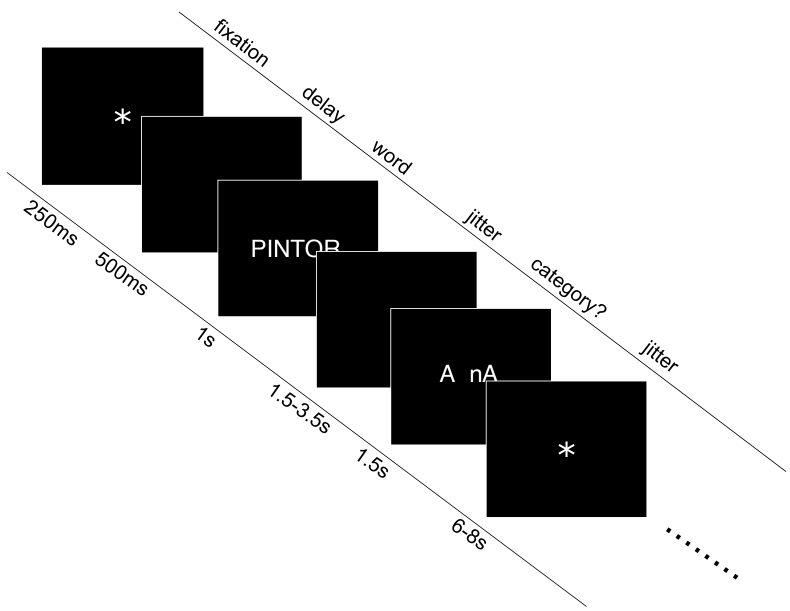


Figure 4.1: The figure illustrates the sequence of events on each trial. Following a fixation period of 250 ms and a delay of 500 ms, a word was presented for 1 s. Next, after a jittered interval of 1.5-3.5 s, participants responded to one question i.e. which category from among animals and non-animals does the word belong to? Finally, the response text ("A nA"/"nA A") was replaced by an asterisk which stayed on the screen for a jittered interval of 6-8 s.

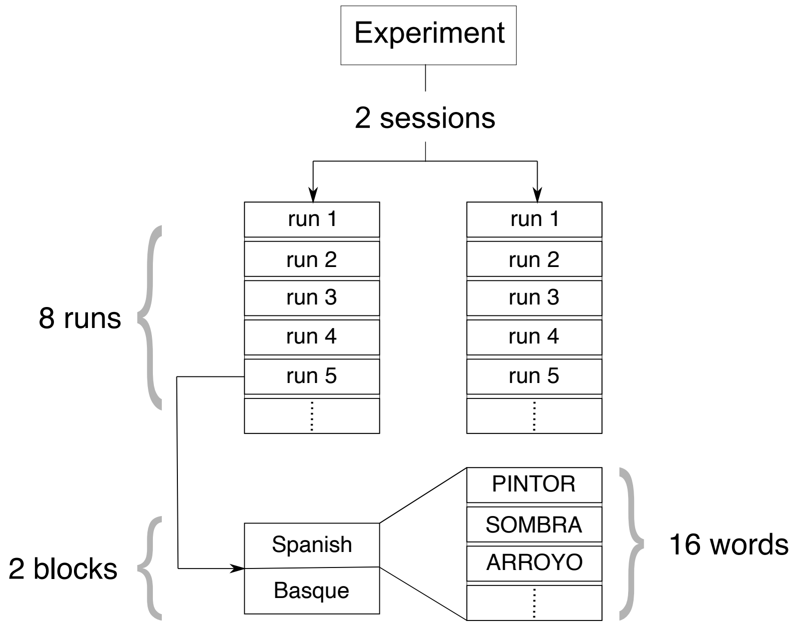


Figure 4.2: The figure summarises the organization of the experiment including sessions, runs, blocks and trials. The experiment comprised of 2 sessions with 8 runs each. Each run was further subdivided into 2 language blocks (Spanish and Basque). Each of these blocks was made up of 16 trials corresponding to single presentation of 16 words randomly drawn from a pool of 128 animal and non-animal words.

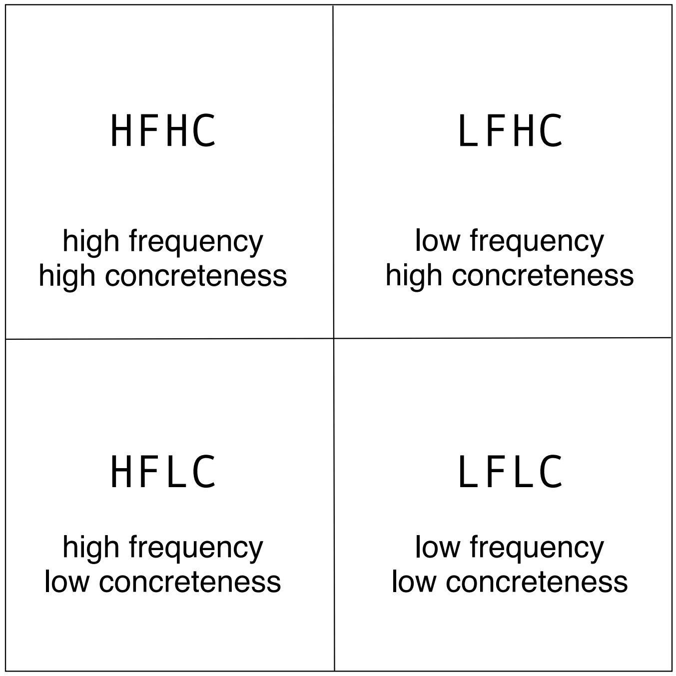


Figure 4.3: The proposed experimental design consisted of two factors (i.e. frequency and concreteness) each with two levels (i.e. low and high) thus creating four different conditions. Correspondingly, the 128-word stimuli were chosen based on the underlying concepts having high/low frequency and high/low concreteness resulting in 32 words per each condition.

To have as many trials as possible per each run, and at the same time maximize the separation between the brain activity corresponding to each of the trials, an event-related design was used and both the stimulus-response and inter-trial gaps were jittered and selected based on pseudo-exponential distributions over 1.5-3.5 s and 6-8 s intervals respectively.

Both instructions and stimuli were presented at the center of the screen, in white against black background and in all uppercase Arial font. The experiment was programmed using Psychopy [123] and is summarized in Figure 4.2. It comprised 2 sessions of 1.5 hours each. These sessions were conducted on the same day or with a difference of 1-7 days depending on the availability of the MRI machine and the participant. Each session was made up of 8 runs of 7 minutes each. Each run in turn was made up of two language blocks i.e. Spanish and Basque blocks each presenting 16 animal and non-animal words.

4.2.5 MRI Data Preprocessing

The preprocessing of fMRI data was performed using FEAT (fMRI Expert Analysis Tool), a tool in FSL suite (FMRIB Software Library; v5.0). After converting all data from DICOM to NIfTI format using MRIConvert (`http://lcni.uoregon.edu/downloads/mriconvert`), the following steps were performed on each run's fMRI. To ensure steady state magnetisation, the first 9 volumes corresponding to the task instruction period were discarded; to remove non-brain tissue, brain extraction tool (BET) [124] was used; head-motion was accounted for using MCFLIRT [125]; minimal spatial smoothing was performed using a gaussian kernel with FWHM of 3 mm. Next, ICA based automatic removal of motion artifacts (ICA-AROMA) was used to remove motion-induced signal variations [163] and this was followed by a high-pass filter with a cutoff of 60 s. All the runs from both first and second session were aligned to a reference volume of the first run (of the first session). All further analyses were performed in native BOLD space.

A set of 15 left-lateralized ROIs was pre-specified (see Figure 4.4) based on a meta-analysis of the semantic system by Binder et al. 2009 [1]. These included: inferior parietal lobe (IPL), inferior temporal lobe (ITL), middle temporal lobe (MTL), fusiform gyrus (FFG), parahippocampal gyrus (PHG), superior frontal gyrus (SFG), posterior cingulate gyrus (PCG), pars opercularis (POP), pars triangularis (PTR), pars orbitalis (POR), frontal pole (FP), medial orbitofrontal cortex (MOFC), laterial orbitofrontal cortex (LOFC), anterior temporal lobe (ATL) and precuneus. First, automatic segmentation of the high-resolution structural image was obtained using FreeSurfer's automated algorithm `recon-all`. Next, `mri_binarize` was used to extract individual gray matter masks from aparc+aseg volume using corresponding label indices in FreeSurferColorLUT text file (`https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/AnatomicalROI`). And finally, after visually inspecting these in FSLView, they were transformed to each run's functional space using FLIRT (7 DoF global rescale transformation). [125, 126] and were binarized (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT/FAQ`).

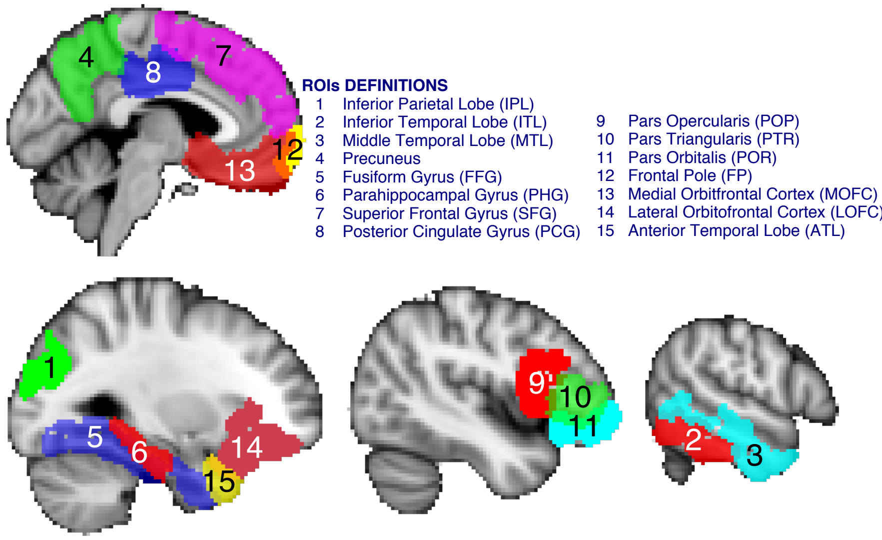


Figure 4.4: The figure shows the selected regions of interest (ROIs) projected on an MNI standard template image. These 15 left-lateralized areas were pre-specified and included inferior parietal lobe (IPL), inferior temporal lobe (ITL), middle temporal lobe (MTL), fusiform gyrus (FFG), parahippocampal gyrus (PHG), superior frontal gyrus (SFG), posterior cingulate gyrus (PCG), pars opercularis (POP), pars triangularis (PTR), pars orbitalis (POR), frontal pole (FP), medial orbitofrontal cortex (MOFC), laterial orbitofrontal cortex (LOFC), anterior temporal lobe (ATL) and precuneus.

4.2.6 Multivariate Pattern Analysis

Multivariate pattern analysis was conducted using scikit-learn [127] and PyMVPA [128] libraries. Classification based on a supervised machine learning algorithm i.e. linear support vector machine [129], was used to evaluate whether multi-voxel patterns in each of the fifteen ROIs carry information related to the semantic category (animal, non-animal) of the word in each of the conditions. Specifically, cross-language (or language-independent) generalization analysis was conducted which entailed training the classifier on trials from one language and testing it on trials from another language. This analysis was done separately for each of the conditions. Additional details related to the data preparation, feature selection, classification and statistics are presented in the following subsections.

4.2.6.1 Data Preparation

For each participant, the relevant time points or scans of the preprocessed fMRI data of each run were labeled with attributes such as word, category, language, and condition using Psychopy generated data files (CSVs). Invariant voxels (or features) were removed. These were the voxels/features whose value did not vary throughout the length of one run. If not removed, such features can cause numerical difficulties with procedures like z-scoring of features. Next, data from all sixteen runs were stacked and each voxel's time series was run-wise z-score normalized and linear detrended. Finally, following two recent cross-language generalization studies [71, 72], one example was created per trial by averaging the 4 volumes between the interval of 3.4 s and 6.8 s after the word onset, which corresponded to 1 second presentation of the word (see Figure 4.1).

4.2.6.2 Pattern Classification

Linear support vector machine (SVM) classifier, with all parameters set to default values as provided by the scikit-learn package ($l2$ regularization, $C = 1.0$, $tolerance = 0.0001$), was used for cross-language decoding in all four conditions. The following procedure was repeated for each ROI separately. To obtain an unbiased generalization estimate, following Varoquaux et al. 2016 [130] the data was randomly shuffled and resampled multiple times to create 300 sets of balanced train-test (80%-20%) splits. Since each example was represented by a single feature vector with each feature a mean of voxel intensities across the sub-interval of 3.4 s and 6.8 s (see Section 4.2.6.1), the length of a vector was equal to the number of voxels in the ROI. To further reduce the dimensionality of the data and thus reduce the chances of overfitting [131, 132], Principal Component Analysis (PCA) with all parameters set to default values as provided by the scikit-learn was used. Since the `n_components` argument was set to `None`, the number of components was chosen to be the smaller from among the number of samples (m) and features (n). In our case, since $n \gg m$, hence, the first m components were selected. The size of the data matrix after PCA was therefore $m \times m$. These components were linear combinations of the preprocessed voxel data and since none of the components was excluded, it was an information loss-less change of the coordinate system to a subspace spanned by the examples [133]. Features thus created were used to train the decoder, and its classification performance on the test set was recorded. This procedure was repeated separately for each of the 300 sets, and the mean of corresponding accuracies was collected for each of the participants. Note that PCA was performed on the training set; then the trained PCA was used to extract components in the test data and its classification performance was assessed. This procedure was repeated separately for each of the 300 sets, and the mean of corresponding accuracies was collected for each of the participants.

4.2.6.3 Statistics

To determine whether the observed generalization accuracy in a given ROI is statistically significantly different from the chance-level of 0.5 (or 50%), a two-tailed t-test was performed. To measure the effect of lexico-semantic factors on the corresponding behavioral identification performance, 2×2 ANOVAs with factors of word frequency and concreteness were conducted. Similarly, to assess the effect of these factors on the corresponding cross-language generalization accuracy, 2×2×15 ANOVAs with factors of word frequency, concreteness and ROIs were conducted. Furthermore, an interaction was expected with high frequency, high concreteness words to lead to a better performance as compared to other three conditions. The p-values corresponding to each of the ROIs were corrected for multiple comparisons using a false discovery rate (FDR) method.

4.3 Results

4.3.1 Behavioral Results

Figure 4.5 presents a summary of the behavioral performance of the participants in the semantic category identification task. Specifically, it shows the percentage of correct responses in each of the four conditions including high frequency-high concreteness (HFHC), high frequency-low concreteness (HFLC), low frequency-high concreteness (LFHC) and low frequency-low concreteness (LFLC). A 2×2 ANOVA showed the main effect of frequency ($F(1, 17) = 2.87, p = 0.11$ for Spanish; $F(1, 17) = 0.007, p = 0.94$ for Basque) to be not statistically significant.

Similarly, it showed the main effect of concreteness ($F(1, 17) = 0.58, p = 0.46$ for Spanish; $F(1, 17) = 2.51, p = 0.13$ for Basque) to be not statistically significant. However, the effect of the interaction between the two factors was found to be statistically significant ($F(1, 17) = 6.32, p < 0.05$ for Spanish; $F(1, 17) = 24.27, p < 0.05$ for Basque).

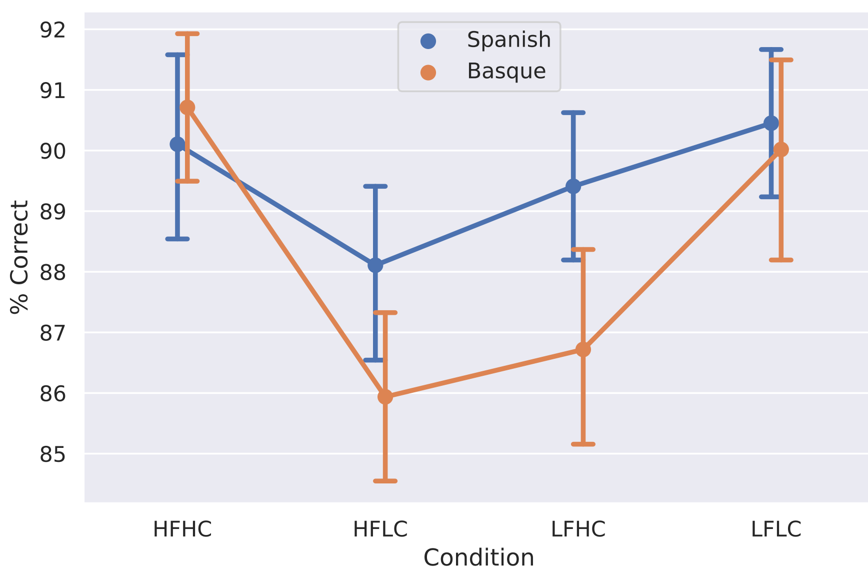


Figure 4.5: The figure presents a summary of the participants' behavioral performance (in percentage) in the semantic category identification task in each of the four conditions. No statistically significant difference in performance was found between the two levels of frequency and concreteness. The p-values were corrected for multiple comparisons.

Figure 4.6, on the other hand, presents summary statistics related to the response time (in seconds) corresponding to the semantic category identification task in each of the four conditions. Specifically, it shows mean and variance of the response time in each of the four conditions including HFHC, HFLC, LFHC and LFLC. A 2×2 ANOVA showed both the main effect of frequency ($F(1, 17) = 2.47, p = 0.13$ for Spanish; $F(1, 17) = 0.42, p = 0.52$ for Basque), concreteness ($F(1, 17) = 3.70, p = 0.07$ for Spanish; $F(1, 17) = 0.36, p = 0.56$ for Basque) and the effect of the interaction between the two factors ($F(1, 17) = 0.54, p = 0.47$ for Spanish; $F(1, 17) = 2.65, p = 0.12$ for Basque) to be not statistically significant.

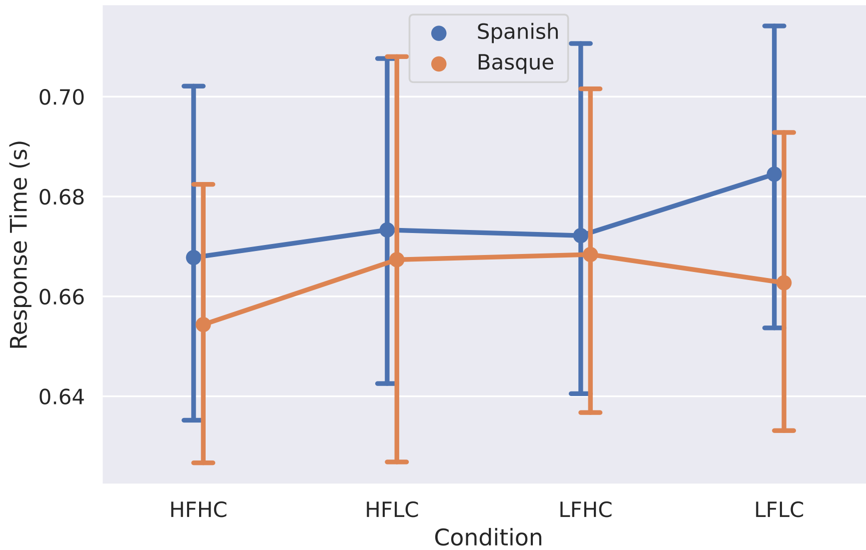


Figure 4.6: The figure presents a summary of the participants' response time (in seconds) for the semantic category identification task in each of the four conditions. No statistically significant difference in performance was found between the two levels of frequency and concreteness. The p-values were corrected for multiple comparisons.

4.3.2 Brain Imaging Results

A 2×2×15 repeated measures ANOVA on the decoding accuracy showed no effect of frequency ($F(1, 17) = 0.24, p = 0.63$) but significant main effects of ROI ($F(14, 238) = 3.36, p < 0.05$) and concreteness ($F(1, 17) = 353.77, p < 0.05$). On the other hand, the interaction between ROI and frequency was statistically significant ($F(14, 238) = 5.34, p < 0.05$) while the effect of the interaction between ROI and concreteness was not ($F(14, 238) = 0.93, p = 0.53$). The effect of the interaction between frequency and concreteness was also not reliable ($F(1, 17) = 1.31, p = 0.27$). Similarly, the effect of the interaction between ROI, frequency and concreteness was also found to be not statistically significant ($F(14, 238) = 1.22, p = 0.26$).

Next, we analysed the cross-language generalization performance. Figures 4.7 to 4.11 present summary statistics related to the cross-language generalization performance in each of the four conditions including HFHC, HFLC, LFHC and LFLC. The generalization performance was found to be statistically significantly above chance in HFHC condition in 9 out of 15 ROIs including frontal pole

($59.27 \pm 4.47\%; t = 8.80; p = 6 \times 10^{-6}$), lateral orbitofrontal cortex ($54.72 \pm 3.58\%; t = 5.60; p = 0.002$), medial orbitofrontal cortex ($55.33 \pm 4.25\%; t = 5.32; p = 0.003$), parahippocampal gyrus ($55.74 \pm 4.43\%; t = 5.50; p = 0.002$), pars orbitalis ($56.77 \pm 4.46\%; t = 6.44; p = 0.0004$), pars triangularis ($55.28 \pm 4.89\%; t = 4.58; p = 0.016$), posterior cingulate gyrus ($57.15 \pm 5.19\%; t = 5.85; p = 0.001$), precuneus ($56.15 \pm 5.32\%; t = 4.91; p = 0.008$) and anterior temporal lobe ($57.06 \pm 4.29\%; t = 6.99; p = 0.0001$). The detailed results can be found in the Appendix B.2.

Similarly, it was found to be statistically significantly above chance in LFHC condition in 7 out of 15 ROIs including lateral orbitofrontal cortex ($56.03 \pm 4.15\%; t = 6.16; p = 0.0006$), medial orbitofrontal cortex ($53.82 \pm 2.84\%; t = 5.70; p = 0.001$), middle temporal lobe ($53.14 \pm 3.24\%; t = 4.11; p = 0.043$), pars orbitalis ($55.25 \pm 5.46\%; t = 4.07; p = 0.047$), pars triangularis ($55.48 \pm 5.35\%; t = 4.34; p = 0.027$), superior frontal gyrus ($54.42 \pm 3.21\%; t = 5.84; p = 0.001$) and anterior temporal lobe ($56.02 \pm 5.35\%; t = 4.77; p = 0.01$). The detailed results can be found in the Appendix B.2.

Interestingly, in HFLC and and LFLC conditions, the cross-language generalization performance was found to be at chance-level ($p > 0.05$) in all 15 pre-specified semantic ROIs (see Appendix B.2).

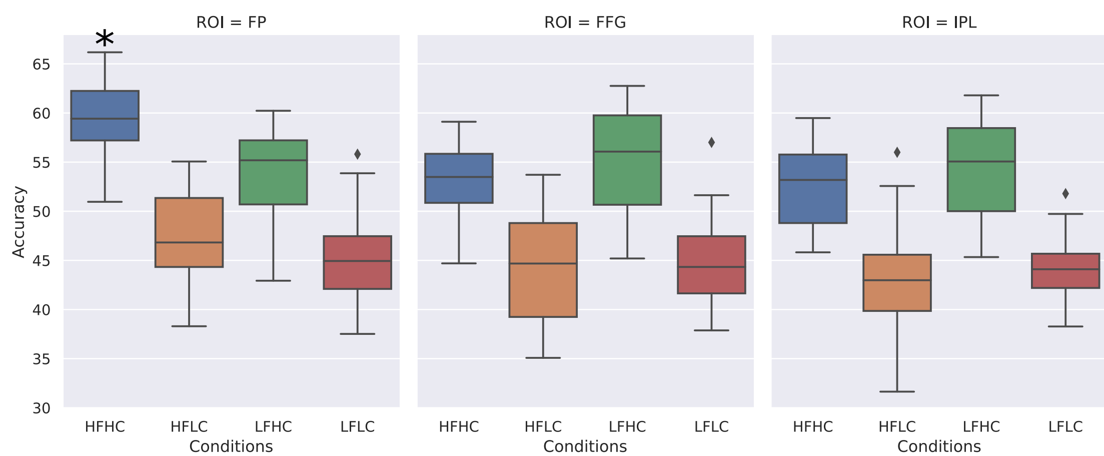


Figure 4.7: The figure shows summary statistics of three of the ROIs including frontal pole, fusiform gyrus and inferior parietal lobe for cross-language generalization in all four conditions. It can be seen that in the frontal pole, the cross-language generalization performance was found to be statistically significantly above chance in HFHC condition. The p-values were corrected for multiple comparisons.

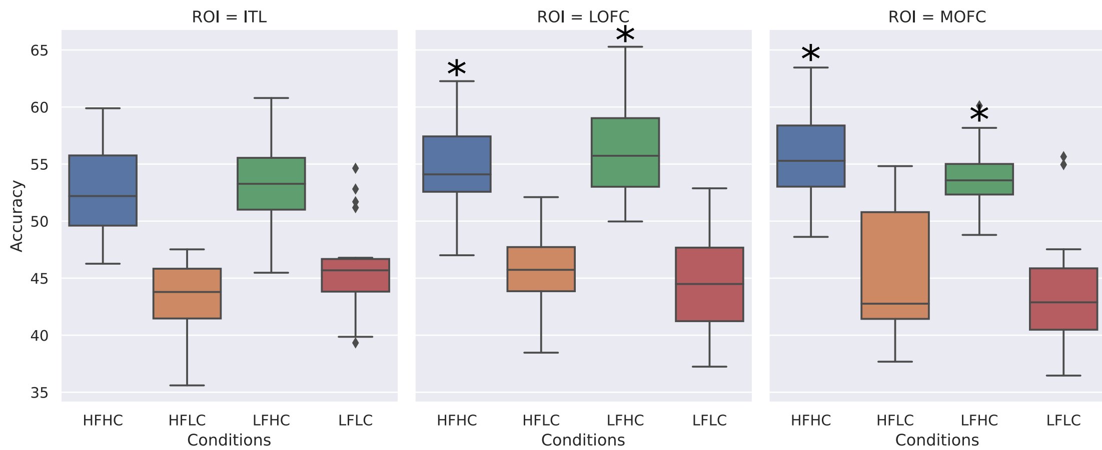


Figure 4.8: The figure shows summary statistics of three of the ROIs including inferior temporal lobe, lateral orbitofrontal cortex and medial orbitofrontal cortex for cross-language generalization in all four conditions. It can be seen that in the lateral and medial orbitofrontal cortex, the cross-language generalization performance was found to be statistically significantly above chance in both HFHC and LFHC conditions. The p-values were corrected for multiple comparisons.

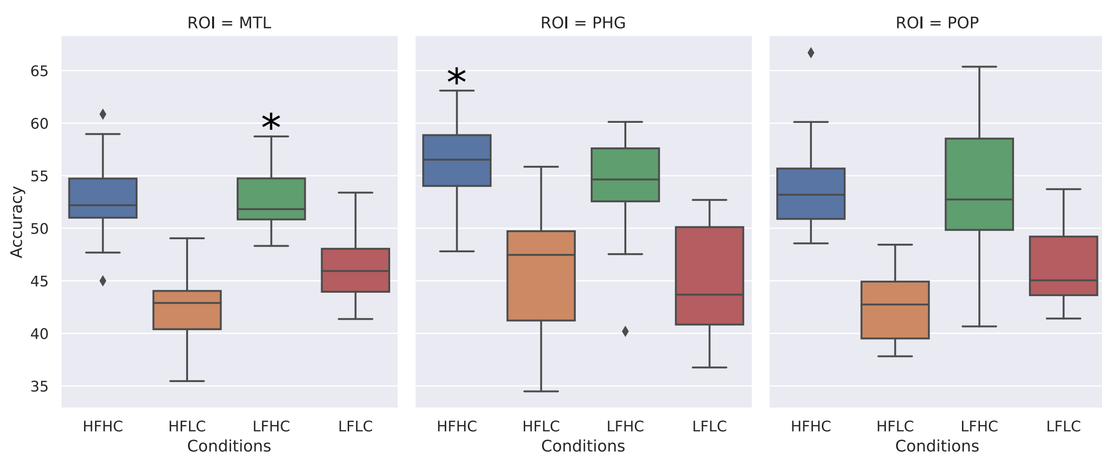


Figure 4.9: The figure shows summary statistics of three of the ROIs including middle temporal lobe, parahippocampal gyrus and pars opercularis for cross-language generalization in all four conditions. It can be seen that in the middle temporal lobe the cross-language generalization performance was found to be statistically significantly above chance in LFHC condition while in the parahippocampal gyrus, it was found to be statistically significantly above chance in HFHC condition. The p-values were corrected for multiple comparisons.

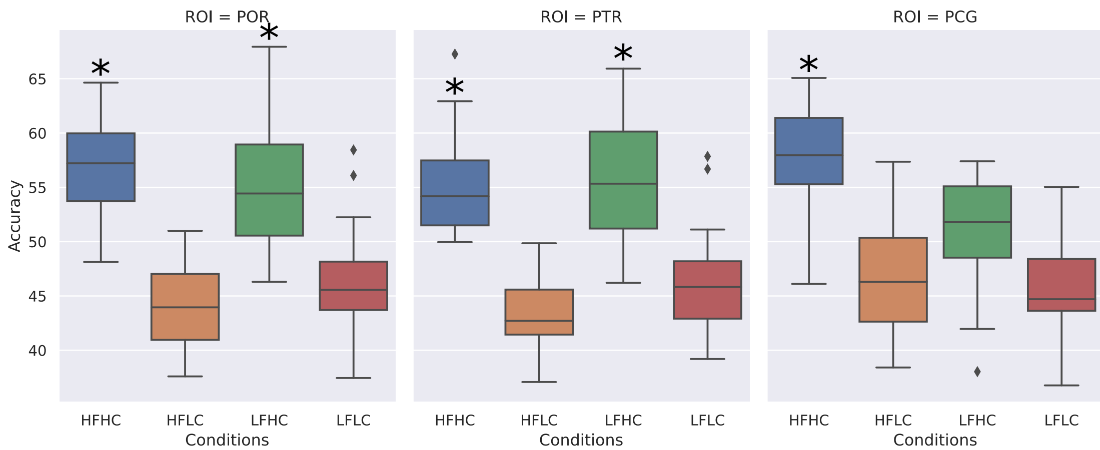


Figure 4.10: The figure shows summary statistics of three of the ROIs including pars orbitalis, pars triangularis and posterior cingulate gyrus for cross-language generalization in all four conditions. It can be seen that in the pars orbitalis and pars triangularis, the cross-language generalization performance was found to be statistically significantly above chance in both HFHC and LFHC conditions. However, in the posterior cingulate gyrus, it was found to be statistically significantly above chance only in HFHC condition. The p-values were corrected for multiple comparisons.

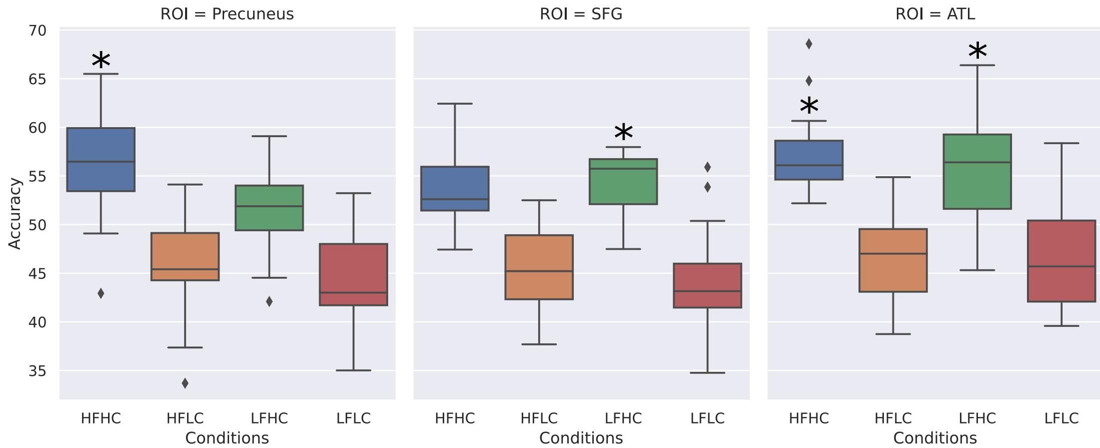


Figure 4.11: The figure shows summary statistics of three of the ROIs including precuneus, superior frontal gyrus and anterior temporal lobe for cross-language generalization in all four conditions. It can be seen that in the precuneus, the cross-language generalization performance was found to be statistically significantly above chance in HFHC condition while in the superior frontal gyrus, it was found to be statistically significantly above chance in LFHC condition. On the other hand, in the anterior temporal lobe, it was found to be statistically significantly above chance in both HFHC and LFHC conditions. The p-values were corrected for multiple comparisons.

4.4 Discussion

In an attempt to address a challenging research question in psychology and neuroscience i.e. whether the word meaning corresponding to different languages of bilinguals is represented within the same neurocognitive system or distributed in separate brain systems, some neuroimaging studies reviewed early in this thesis, used multivariate pattern decoding of corresponding neural activity. Specifically, they adopted a technique referred to as cross-language decoding. The present fMRI study used a diverse set of words with different levels of frequency and concreteness and showed that BOLD activity patterns associated with high frequency, high concreteness words (HFHC) contain information that allows for above-chance cross-language generalization of the semantic category from Spanish to Basque and vice versa in 9 out of 15 ROIs in the semantic network. These regions were found to be distributed across frontal, temporal and parietal lobes. It also showed that the BOLD activity patterns associated with low frequency, high concreteness (LFHC) words allow for above-chance cross-language generalization of the semantic category in 7 out of 15 ROIs. On the other hand, BOLD activity patterns associated with high frequency, low concreteness (HFLC) and low frequency, low concreteness (LFLC) only lead to chance-level cross-language generalization performance in all 15 ROIs.

These results showed word concreteness to have an influence on the language-shared semantic representations. Specifically, words, high in concreteness (HC), lead to above-chance cross-language decoding in FP, LOFC, MOFC, MTL, PHG, POR, PTR, PCG, Precuneus and ATL; while words, low in concreteness (LC), irrespective of whether they were high/low in frequency, lead to chance-level cross-language generalization in all ROIs. This is in line with results of previous chapters where only HC words were used, and above-chance cross-language generalization was found during conscious and deep processing in ROIs including IPL, MTL, ITL, FFG, PHG, IFG, and PCG (see Sections 2.3.2 and 3.3.2). This is also in keeping with previous neuroimaging studies [71, 72], though here we revealed the critical role of lexico-semantic characteristics of stimuli.

If we look at the ROIs with above-chance cross-language generalization, while some of them were found to be common between HFHC and LFHC words including LOFC, MOFC, POR, PTR, and ATL, others were found to be specific to HFHC words including FP, PHG, PCG and Precuneus. Similarly, one ROI was found to be specific to LFHC words i.e. MTL. These observations suggested language-shared semantic representations corresponding to HFHC and LFHC words to be partially overlapping. More research is however needed to characterize the similarities and differences between these representations.

On the other hand, HFLC and LFLC words lead to only chance-level cross-language generalization in all 15 ROIs. In other words, HF words, lead to above-chance cross-language generalization if and only if they were also high in concreteness (HFHC). It can be argued that the proposed ROI-based multivariate analysis may not be well-suited to localize the language-shared semantic representations of LC words.

The influential hub-and-spoke model suggests that unified and amodal representations are formed within a single hub i.e. anterior temporal lobes (ATL). In our study, we found significant cross-language semantic generalization of HFHC and LFHC words in a distributed set of semantic ROIs including ATL, but only chance-level cross-language generalization was observed for HFLC and LFLC words even in ATL. These null results however must be taken with caution given that ATL is well-known to have susceptibility-induced signal dropout issues, and also considering the amount of evidence in the favour of the key role of ATL as a multi-modal semantic hub [173].

These findings also have implications for psycholinguistic models of visual word recognition such as BIA+ [56]. Our results indicate that non-selective access to word meaning across languages is not an intrinsic property of the semantic system. Instead, they are in keeping with the view that depending on the level of concreteness and frequency of the concepts, the extent of parallel and non-selective access can be modulated. More research is however needed to elucidate the extent to which lexico-semantic factors shape bilinguals' access to semantic representations, namely, the extent to which different language representations are co-activated in parallel [180].

Chapter 5

General Discussion

The main goal of this dissertation was to investigate the neural underpinnings of semantic processing in bilinguals. Specifically, to investigate the factors that influence the cross-language generalization of semantic representations from L1 to L2 and vice versa. Among the main factors considered were: state of visual awareness, depth of word processing and lexico-semantic characteristics of stimuli. Using fMRI-based MVPA to achieve this goal allowed answering important research questions including: whether brain can encode the meaning of words in the absence of conscious awareness (see chapter 2)? whether different languages in bilinguals are integrated in the same system with overlapping semantic representations or rely on separate representations for each language (see chapter 3)? whether semantic representations corresponding to abstract and relatively less frequent words also generalize across languages (see chapter 4)? and how are semantic representations corresponding to abstract and less frequent words similar to and different from those of concrete and frequent words (see chapter 4)?

In this final discussion chapter, the most significant empirical findings made are summarized relating them with existing theories of meaning representation and psycholinguistic models. Finally, the chapter is concluded with some limitations of the studies conducted and discussion of future prospects.

5.1 A Summary of Empirical Findings

Previous neuroimaging studies of bilinguals showed that the brain activity patterns created during semantic processing in L1 generalize to those created during semantic processing in L2 ([71, 72]). Specifically, these fMRI studies showed certain areas of the brain (e.g. left parietal lobe, inferior frontal gyrus and posterior temporal lobe) to have above-chance cross-language generalization, suggesting the presence of language-shared semantic representations. A key limitation of these studies however was that they assumed the activation of language-shared semantic representations to be automatic and therefore did not consider potential factors that may mediate the generalization of semantic representations across languages.

The Chapter 2 of this thesis presented an fMRI-based MVPA study of Spanish-Basque bilinguals, which used masked animal/non-animal words as stimuli, and investigated if the level of visual awareness influences the cross-language generalization of semantic representations from Spanish to Basque and vice versa. A brief discussion of the most important empirical findings of this study are presented as follows.

Recall that, in this study, the level of awareness of each words was assessed through a subjective rating scale and objective performance on the task of animal/non-animal categorization. These criteria were then used to categorize the trials as fully conscious, partially conscious or non-conscious. In partially and fully conscious conditions, all seven canonical areas of the semantic network were found to contain the semantic representation of word category. On the other hand, in the non-conscious condition, four ROIs (IPL, dmPFC, IFG, and PCG) were implicated for Spanish, and two (VTL, PCG) for Basque. Importantly, cross-language generalization was found to be at chance-level for both Spanish to Basque and vice versa. This happened also for partially conscious trials, and only some evidence of cross-language generalization was found in the fully conscious condition. This indicates that conscious awareness may be a critical factor for cross-language generalization and that this is not mandatory for non-conscious words.

The chapter 3 of this thesis provided novel insights into how the depth of processing during semantic tasks influences cross-language generalization based on fMRI-based MVPA in putative substrates of the semantic network. We found that the cross-language generalization was not different from chance in all ROIs during shallow processing conditions i.e. when participants were asked to merely attend and read the words. Only in the context of deep information processing i.e. when participants were instructed to read and think about the meaning, did brain activity patterns reliably generalize in several brain regions (from Spanish to Basque: IPL, LTL, VTL, dmPFC, IFG, and PCG; from Basque to Spanish: IPL, LTL, VTL and IFG). Importantly, significant cross-language generalization was observed in inferior frontal cortex, previously known to be involved in semantic control e.g. switching between two languages of bilinguals [175, 176, 177].

The chapter 4 of this thesis used a diverse set of words with different levels of frequency (low, high) and concreteness (low, high) and showed that BOLD activity patterns associated with high frequency, high concreteness words (HFHC) contain information that allows for above-chance cross-language generalization of the semantic category from Spanish to Basque and vice versa in 9 out of 15 pre-specified ROIs of the semantic network. These regions were found to be distributed across frontal, temporal and parietal lobes. It also showed that the BOLD activity patterns associated with low frequency, high concreteness (LFHC) words allowed for above-chance cross-language generalization of the semantic category in 7 out of 15 ROIs. While 5 of these 7 ROIs were found to be common between HFHC and LFHC words, superior frontal lobe and middle temporal lobe were found to be specific to cross-language generalization of LFHC words.

5.2 Theoretical Implications

The outcomes of this thesis have ramifications for psycholinguistic models of visual word recognition (see section 1.2.2) and theories of meaning representation e.g. hybrid/pluralistic accounts of meaning representation (see section 1.1.3).

Psycholinguistic models like BIA+ [56] implement word processing in a purely bottom-up manner with parallel and non-selective (i.e. language independent) activation of linguistic codes not just at the level of semantics but orthography and phonology too. We propose that such models need to be revised to incorporate the influence of top-down factors. Our results indicated that non-selective access to word meaning across languages is not mandatory or intrinsic property of the semantic system. Instead, they are in keeping with the view that depending on factors including level of awareness, and depth of processing, the extent of parallel and non-selective access can be modulated. More research is however needed to elucidate the extent to which these factors shape how bilinguals access semantic representations, namely, the extent to which different language representations for a given concept are co-activated in parallel [180] or whether, according to BIA+, bilinguals access to the lexical and semantic representation is delayed in the second language compared to the first language [181].

Recall that hybrid theories like hub-and-spoke model suggest that sensory-motor representations of a concept are encoded in modality-specific brain regions, yet, unified and amodal representations are formed within a single transmodal hub such as anterior temporal lobes (ATL). Importantly, in this thesis, we found the cross-language generalization in multiple substrates of the semantic network including ATL. Consequently, we propose that the conscious, deep information processing of concrete and frequent words triggered the global sharing of information across a distributed set of brain areas implicated in semantic representation and this supported cross-language generalization. We propose that our results are in keeping with distributed-only views of semantic processing [15].

5.3 Limitations and Future Prospects

There are a number of reasons behind not finding a strong evidence of cross-language generalization even in fully conscious condition in Chapter 2. First, the experiment was designed to maximize the number of non-conscious trials. The stimuli was briefly presented and masked, and luminance was varied based on a staircase procedure that was biased towards decreasing luminance. Accordingly, the experimental task may only have promoted shallow encoding of the words. Given the relatively small number of words used, it is also possible that the observers learned a mapping between the low-level properties of the word stimuli and the semantic categorization response, which did not involve the level of processing required for across language generalization. We suggested that our task may have promoted a level of processing that is insufficient for cross-language generalization.

Another important limitation is that, for multivariate pattern analysis, ROI-based approach was used. It was chosen as a middle ground between searchlight analysis [193] which leverages the correlation structure among a local set of voxels, and sparse multinomial linear regression (SMLR) [166] which takes into account potential correlations between spatially remote voxels. It can be argued that this approach may not be well-suited to measure the effect of all the factors investigated. For example, it may be the case that abstract words (LC words; see Chapter 4) activate language-shared representations with corresponding informative patterns either being too localized or too spatially removed to be captured using proposed ROI-based analysis. Future research should use ROI-based analysis together with searchlight and SMLR approaches to provide a more complete picture. However, we believe that our ROI based analyses are sensitive enough in that robust decoding was observed across the studies.

5.4 Conclusion

The three empirical studies presented in the current dissertation contributed to research on bilingual semantic representations. Overall, it can be concluded that such factors as the state of visual awareness (Chapter 2), the depth of word processing (Chapter 3) and the lexico-semantic characteristics of stimuli (Chapter 4) influence the generalization of semantic representations across languages.

Resumen Amplio en Castellano

Cuando se trata del procesamiento semántico en bilingües, una pregunta importante es si los diferentes lenguajes están integrados en el mismo sistema con representaciones compartidas o se basan en representaciones separadas para cada lenguaje. La evidencia conductual sugiere que las representaciones semánticas se superponen, al menos parcialmente. Esto ha motivado el desarrollo de modelos psicolingüísticos de representación del lenguaje bilingüe. Aunque estos modelos difieren en sus predicciones sobre los mecanismos que subyacen al procesamiento léxico y los vínculos entre el procesamiento léxico y semántico de los dos lenguajes, coinciden en que las representaciones semánticas se superponen, al menos parcialmente, entre los lenguajes. Los estudios de imágenes cerebrales, basados en imágenes de resonancia magnética funcional (IRMf) y análisis univariados, mostraron cierta evidencia de respuestas cerebrales tanto compartidas por el lenguaje como específicas del lenguaje, pero no se puede descartar el papel de factores estratégicos como las listas de expectativas de relaciones prime-target. Los enfoques univariados, por otro lado, no son los más adecuados para identificar si el procesamiento semántico está mediado o no por un sistema similar en diferentes lenguajes. La observación de que un área cortical se activa en ambos lenguajes no implica que las representaciones cerebrales también sean similares. Estudios recientes utilizaron el análisis de patrones multivariados (APMV) para evaluar si los patrones de actividad cerebral provocados por palabras en un lenguaje pueden predecir los patrones de palabras equivalentes en el otro lenguaje. Los resultados mostraron la existencia de representaciones compartidas por el lenguaje en sustratos semánticos bien conocidos, incluyendo el lóbulo parietal izquierdo, la circunvolución frontal inferior y el lóbulo temporal posterior. Una limitación clave de estos estudios es que queda por determinar los factores que subyacen a la generalización de las representaciones semánticas entre los lenguajes. Es importante destacar que ninguno de estos estudios consideró factores tales como la conciencia visual, la profundidad del procesamiento y las propiedades léxico-semánticas de los estímulos.

El objetivo central de esta disertación fue investigar cómo factores que incluyen el estado de la conciencia visual, la profundidad del procesamiento de palabras y las características léxico-semánticas de las palabras influyen en la generalización de las representaciones semánticas entre lenguajes.

En el primer capítulo empírico, se consideró el factor de la conciencia visual. Específicamente, se utilizó un estudio APMV basado en IRMf de bilingües Español-Euskera, utilizando equivalentes traduccionales de Español y Euskera animales/no animales como estímulos, para investigar si la categoría semántica (animal/no animal) de las palabras parcialmente consciente o no consciente se pueden decodificar a partir de patrones de actividad de múltiples vóxeles en áreas semánticas putativas del cerebro. En segundo lugar, utilizando la decodificación entre lenguajes, se observó cómo los diferentes niveles de conciencia (total, parcialmente y no consciente) afectan la generalización de la categoría semántica entre lenguajes en las áreas semánticas bien conocidas del cerebro. Debido a que los niveles de conciencia se controlaron mediante calificaciones subjetivas, así como medidas de desempeño objetivas, y los estímulos en Español/Euskera se controlaron según las propiedades lingüísticas (por ejemplo, longitud, frecuencia y afines, etc.) en todos los lenguajes y categorías, este estudio ofreció pruebas sólidas relacionado con la influencia de los niveles de conciencia sobre la representación semántica en general y las representaciones lingüísticas compartidas en particular. Específicamente, este estudio demostró que el significado de las palabras no conscientes se puede codificar en patrones de actividad de múltiples vóxeles en regiones semánticas putativas, incluidas las áreas frontales. Mientras que la clasificación del significado de las palabras dentro del lenguaje era posible en contextos no conscientes, la generalización de significado entre lenguajes requirió un análisis semántico consciente y más profundo.

En el segundo capítulo empírico, se consideró el factor de profundidad de procesamiento. Específicamente, se utilizó un estudio APMV basado en IRMf de bilingües, utilizando equivalentes de traducción en Español y Euskera como estímulos, para investigar cómo la profundidad del procesamiento de palabras (superficial o profunda) afecta la decodificación de la categoría semántica de patrones de actividad de múltiples vóxeles en áreas semánticas putativas del cerebro. La profundidad del procesamiento se varió al motivar un procesamiento superficial, es decir, solo leer en la mitad de las pruebas y un procesamiento relativamente más profundo, es decir, leer acompañado con pensar en el significado en la otra mitad de las pruebas. Utilizando la decodificación entre lenguajes, se investigó cómo la profundidad del procesamiento media la generalización de la categoría semántica entre lenguajes en áreas semánticas bien conocidas del cerebro. Descubrimos que la generalización entre lenguajes no era diferente del azar en todas las regiones de interés (RIs) durante la condición de procesamiento superficial. Solo en el contexto del procesamiento profundo de la información, es decir, cuando se instruyó a los participantes para que leyeran y pensaran en el significado, los patrones de actividad cerebral se generalizaron de manera confiable en varios RIs (del Español al Euskera: lóbulo parietal inferior, lóbulo temporal lateral, lóbulo temporal ventromedial, corteza prefrontal dorsomedial, giro frontal inferior y giro cingulado posterior; del Euskera al Español: lóbulo parietal inferior, lóbulo temporal lateral, lóbulo temporal ventromedial y lóbulo frontal inferior). Es importante destacar que se observó una generalización significativa entre lenguajes en áreas delfrontal inferior, que anteriormente se sabía que estaba involucrada en el control semántico, por ejemplo, controlar cambio entre dos lenguajes.

En el tercer capítulo empírico, se consideraron algunos factores léxico-semánticos que incluyen la frecuencia y la concreción de las palabras. Específicamente, se utilizó un estudio APMV basado en IRMf de bilingües, que utilizó un conjunto extenso de equivalentes traduccionales en Español y Euskera animales/no animales con diferentes niveles de concreción y frecuencia como estímulos, para investigar cómo los factores léxico-semánticos influyen en la generalización de la categoría semántica entre lenguajes en sustratos canónicos de la red semántica. El hecho de que este estudio utilizó un conjunto de estímulos comparativamente más grande le dio una ventaja sobre otros dos estudios en cuanto a los tipos de análisis multivariante que se pueden realizar y las inferencias generales que se pueden extraer. Los resultados mostraron que la concreción influye en las representaciones semánticas compartidas por el lenguaje. Específicamente, las palabras, con un alto contenido de concreción, conducen a una decodificación del significado entre lenguajes por encima del azar en el polo frontal, la corteza orbitofrontal lateral, la corteza orbitofrontal medial, el lóbulo temporal medial, la circunvolución parahipocampal, la pars opercularis, la pars triangularis, la circunvolución cingulada posterior, el precúneo y el lóbulo temporal anterior mientras que las palabras, de baja concreción, independientemente de si eran de alta o baja frecuencia, conducen a una generalización entre lenguajes a nivel de probabilidad en todos los RIs. Por otro lado, las palabras de poca concreción, independientemente de si tenían una frecuencia alta o baja, solo conducen a una generalización entre lenguajes a nivel de probabilidad en los 15 RIs. Las palabras de alta frecuencia conducen a una generalización entre lenguajes por encima del azar si y sólo si también tienen un alto grado de concreción. Esto está de acuerdo con los resultados de otros dos estudios empíricos en los que se usaron palabras concretas y familiares, y se encontró una generalización entre lenguajes superior al azar durante el procesamiento consciente y profundo en RIs, incluyendo el lóbulo parietal inferior, el lóbulo temporal medial, el lóbulo temporal inferior, giro fusiforme y parahipocampal, circunvolución frontal inferior y circunvolución cingulada posterior. Esto también está en consonancia con los resultados de estudios previos de neuroimagen bilingüe. Los resultados de estos estudios empíricos tienen ramificaciones para los modelos psicolingüísticos de reconocimiento visual de palabras y teorías de representación de significado, por ejemplo, teorías híbridas/pluralistas de la representación del significado.

Los modelos psicolingüísticos como BIA+ implementan el procesamiento de palabras de una manera puramente ascendente (bottom-up) con activación paralela y no selectiva (es decir, independiente del lenguaje) de códigos lingüísticos no solo a nivel de semántica sino también de ortografía y fonología. Proponemos que dichos modelos deben revisarse para incorporar la influencia de los factores de arriba hacia abajo (top-down). Nuestros resultados indicaron que el acceso no selectivo al significado de las palabras en todos los lenguajes no es una propiedad obligatoria o intrínseca del sistema semántico. En cambio, están de acuerdo con un modelo de que dependiendo de factores que incluyen el nivel de conciencia, la profundidad del procesamiento y las propiedades léxico-semánticas de los estímulos (frecuencia y concreción de los conceptos), se puede modular el alcance del acceso paralelo y no selectivo. Sin embargo, se necesita más investigación para dilucidar hasta qué punto estos factores dan forma a cómo los bilingües acceden a las representaciones semánticas, es decir, hasta qué punto se co-activan en paralelo diferentes representaciones lingüísticas para un concepto dado. Por otro lado, entre las teorías de la representación del significado, las teorías híbridas como el modelo hub-and-spoke sugieren que las representaciones sensoriomotoras de un concepto están codificadas en regiones del cerebro específicas de la modalidad, sin embargo, las representaciones unificadas y amodales se forman dentro de un centro transmodal único como los lóbulos temporales anteriores (LTA). Es importante destacar que, en esta disertación, encontramos la generalización entre lenguajes en múltiples sustratos de la red semántica, incluyendo LTA. En consecuencia, proponemos que el procesamiento consciente y profundo de la información de palabras concretas y frecuentes desencadenó el intercambio global de información a través de un conjunto distribuido de áreas cerebrales implicadas en la representación semántica y esto apoyó la generalización entre lenguajes. Sugerimos que nuestros resultados están de acuerdo con las teorías distribuidas del procesamiento semántico.

Los tres estudios empíricos presentados en la tesis actual contribuyeron a mejorar nuestra comprensión de las representaciones semánticas bilingües. En general, se puede concluir que factores como el estado de la conciencia visual (Capítulo 2), la profundidad del procesamiento de palabras (Capítulo 3) y las características léxico-semánticas de los estímulos (Capítulo 4) influyen en la generalización de las representaciones semánticas entre diferentes lenguajes.

Appendix A

Supplemental Materials for Chapter 3

A.1 Within-Language Decoding

ROI	shallow		deep		deep - shallow
IPL	58.61±4.40	$p = 3.97 \times 10^{-11}$	62.98±5.85	$p = 8.19 \times 10^{-12}$	$p = 9.77 \times 10^{-3}$
LTL	59.45±4.54	$p = 1.27 \times 10^{-11}$	60.81±6.97	$p = 5.35 \times 10^{-9}$	$p = 0.45$
VTL	58.89±5.19	$p = 5.20 \times 10^{-10}$	61.62±5.57	$p = 1.21 \times 10^{-11}$	$p = 4.77 \times 10^{-2}$
dmPFC	58.80±4.89	$p = 5.40 \times 10^{-11}$	60.18±5.31	$p = 6.39 \times 10^{-11}$	$p = 0.44$
IFG	59.30±4.01	$p = 2.84 \times 10^{-12}$	61.21±7.38	$p = 6.69 \times 10^{-9}$	$p = 0.38$
vmPFC	55.01±5.64	$p = 4.59 \times 10^{-5}$	56.19±5.92	$p = 5.00 \times 10^{-6}$	$p = 0.44$
PCG	58.64±4.10	$p = 1.27 \times 10^{-11}$	62.84±6.07	$p = 1.21 \times 10^{-11}$	$p = 9.24 \times 10^{-3}$
ATL	56.57±6.29	$p = 5.05 \times 10^{-6}$	52.00±5.47	$p = 0.06$	$p = 9.24 \times 10^{-3}$

Table A.1: The table presents within-language decoding results for Spanish in both shallow and deep processing conditions. The p-values were corrected for multiple comparisons.

ROI	shallow		deep		deep - shallow
IPL	58.95±4.18	$p = 1.30\times10^{-11}$	62.62±4.65	$p = 5.28\times10^{-14}$	$p = 0.01$
LTL	58.34±4.75	$p = 4.67\times10^{-10}$	60.44±5.86	$p = 2.29\times10^{-10}$	$p = 0.26$
VTL	58.50±5.27	$p = 1.95\times10^{-9}$	62.73±6.09	$p = 1.13\times10^{-11}$	$p = 0.03$
dmPFC	58.29±4.82	$p = 6.02\times10^{-10}$	60.56±5.12	$p = 1.16\times10^{-11}$	$p = 0.15$
IFG	58.45±4.00	$p = 1.30\times10^{-11}$	61.57±6.06	$p = 5.65\times10^{-11}$	$p = 0.11$
vmPFC	54.80±5.87	$p = 1.51\times10^{-4}$	55.54±5.12	$p = 2.88\times10^{-6}$	$p = 0.69$
PCG	59.12±4.93	$p = 1.94\times10^{-10}$	60.78±4.48	$p = 5.57\times10^{-13}$	$p = 0.30$
ATL	54.15±5.88	$p = 6.85\times10^{-4}$	53.71±5.63	$p = 0.001$	$p = 0.75$

Table A.2: The table presents within-language decoding results for Basque in both shallow and deep processing conditions. The p-values were corrected for multiple comparisons.

A.2 Cross-Language Generalization

ROI	shallow		deep		deep - shallow
IPL	50.62±4.74	p = 0.57	55.18±5.27	$p = 2.99\times10^{-5}$	p = 0.03
LTL	51.04±4.08	p = 0.48	55.84±5.35	$p = 1.78\times10^{-5}$	p = 0.02
VTL	52.22±4.21	p = 0.07	55.45±5.49	$p = 2.99\times10^{-5}$	$p = 4.94\times10^{-2}$
dmPFC	50.85±4.42	p = 0.57	53.12±4.25	$p = 6.00\times10^{-4}$	p = 0.07
IFG	50.47±3.67	p = 0.57	54.89±5.33	$p = 6.00\times10^{-5}$	p = 0.03
vmPFC	50.62±4.55	p = 0.57	51.47±2.74	$p = 8.00\times10^{-3}$	p = 0.92
PCG	50.45±4.59	p = 0.60	53.57±4.62	$p = 4.00\times10^{-4}$	p = 0.06
ATL	51.20±3.72	p = 0.37	50.01±4.69	p = 0.99	p = 0.92

Table A.3: The table presents cross-language generalization results for Spanish to Basque generalization in both shallow and deep processing conditions. The p-values were corrected for multiple comparisons.

ROI	shallow		deep		deep - shallow
IPL	51.66±4.97	p = 0.16	54.50±5.12	$p = 2.00\times10^{-4}$	p = 0.03
LTL	51.08±4.20	p = 0.28	54.47±5.72	$p = 5.00\times10^{-4}$	p = 0.06
VTL	52.36±4.35	p = 0.05	55.34±6.36	$p = 3.00\times10^{-4}$	p = 0.08
dmPFC	50.87±4.00	p = 0.34	53.05±4.38	$p = 1.00\times10^{-3}$	p = 0.12
IFG	51.48±3.18	p = 0.07	54.78±5.06	$p = 2.00\times10^{-4}$	p = 0.03
vmPFC	51.73±4.73	p = 0.16	50.55±3.65	p = 0.48	p = 0.96
PCG	50.67±4.73	p = 0.51	53.03±5.95	p = 0.01	p = 0.14
ATL	50.12±3.99	p = 0.88	49.64±4.28	p = 0.65	p = 0.13

Table A.4: The table presents cross-language generalization results for Basque to Spanish generalization in both shallow and deep processing conditions. The p-values were corrected for multiple comparisons.

A.3 Out of Sample Generalization

Figure A.1 and A.2 present the summary statistics of the ROIs for out-of-sample generalization in both shallow and deep processing conditions. It can be seen that in the shallow processing condition, the decoding of the semantic category (living/non-living) in Spanish was found to be above-chance (FDR corrected for multiple comparisons) in two out of eight ROIs including IPL (51.07±3.96; $t(30) = 1.46$; $p = 0.27$), LTL (52.46±4.35; $t(30) = 3.05$; $p = 0.02$), VTL (51.31±5.00; $t(30) = 1.41$; $p = 0.27$), dmPFC (51.50±4.61; $t(30) = 1.76$; $p = 0.24$), IFG (52.17±3.87; $t(30) = 3.01$; $p = 0.02$), vmPFC (50.29±4.64; $t(30) = 0.34$; $p = 0.74$), PCG (50.75±4.22; $t(30) = 0.96$; $p = 0.39$), ATL (51.31±5.47; $t(30) = 1.29$; $p = 0.28$). In Basque however, it was found to be at chance-level in all pre-specified ROIs including IPL (51.06±4.53; $t(30) = 1.26$; $p = 0.82$), LTL (50.57±4.85; $t(30) = 0.63$; $p = 0.82$), VTL (50.21±4.78; $t(30) = 0.23$; $p = 0.82$), dmPFC (50.23±4.29; $t(30) = 0.29$; $p = 0.82$), IFG (50.34±4.68; $t(30) = 0.39$; $p = 0.82$), vmPFC (49.66±5.29; $t(30) = -0.34$; $p = 0.82$), PCG (51.48±4.50; $t(30) = 1.77$; $p = 0.69$), ATL (49.44±5.42; $t(30) = -0.56$; $p = 0.82$).

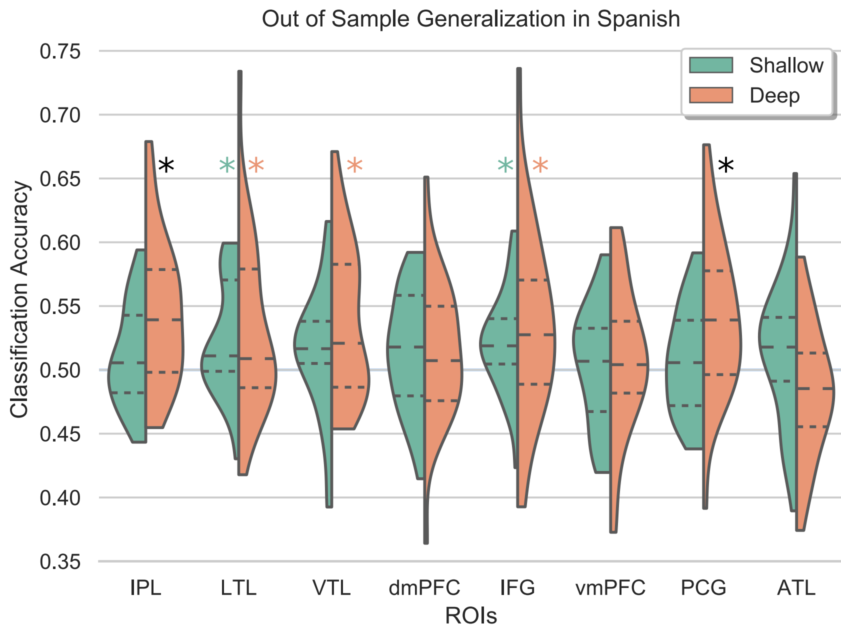


Figure A.1: The figure shows summary statistics of the ROIs for out-of-sample generalization of the semantic category in Spanish. The three dotted lines inside each violin are the quartiles. The green and orange asterisks mark the ROIs that showed significantly above-chance performance in the shallow and deep conditions respectively and the black asterisks those with statistically significant improvement in deep as compared to shallow condition. The p-values were corrected for multiple comparisons.

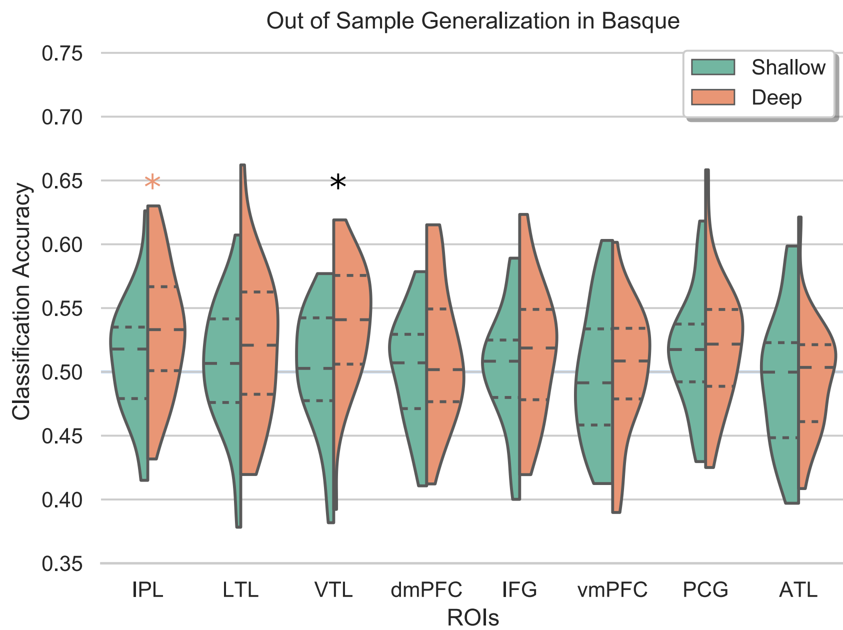


Figure A.2: The figure shows summary statistics of the ROIs for out-of-sample generalization of the semantic category in Basque. The three dotted lines inside each violin are the quartiles. The orange asterisks mark those that showed above-chance performance in the deep condition and the black asterisks mark those with statistically significant improvement in deep as compared to shallow condition. The p-values were corrected for multiple comparisons.

On the other hand, in the deep processing condition, the decoding of the semantic category in Spanish was found to be above-chance and better than shallow condition (FDR corrected for multiple comparisons) in three out of eight ROIs including: IPL (54.39±5.57; $t(30) = 4.24$; $p = 0.002$), LTL (53.05±6.55; $t(30) = 2.51$; $p = 0.029$), VTL (53.77±5.80; $t(30) = 3.49$; $p = 0.004$), dmPFC (51.37±5.43; $t(30) = 1.36$; $p = 0.21$), IFG (53.37±7.22; $t(30) = 2.51$; $p = 0.03$), vmPFC (50.70±5.72; $t(30) = 0.66$; $p = 0.51$), PCG (53.95±5.98; $t(30) = 3.56$; $p = 0.004$), ATL (48.29±5.11; $t(30) = -1.80$; $p = 0.11$). In Basque however, it was found to be above-chance and better than shallow condition (FDR corrected for multiple comparisons) in one out of eight ROIs including: IPL (53.28±5.03; $t(30) = 3.52$; $p = 0.006$), LTL (51.87±5.73; $t(30) = 1.76$; $p = 0.14$), VTL (53.69±4.81; $t(30) = 4.13$; $p = 0.002$), dmPFC (51.45±5.33; $t(30) = 1.47$; $p = 0.20$), IFG (51.68±5.14; $t(30) = 1.76$; $p = 0.14$), vmPFC (50.11±4.84; $t(30) = 0.12$; $p = 0.90$), PCG (52.11±4.82; $t(30) =$

2.36; $p = 0.07$), ATL (49.53 ± 4.23; $t(30) = -0.59$; $p = 0.64$).

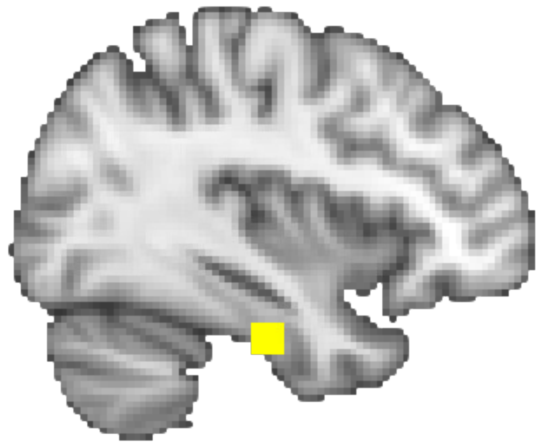


Figure A.3: The figure shows another location of anterior temporal lobe projected on an MNI standard template image. It is a more posterior area previously implicated as a semantic hub by [2].

A.4 Bayesian Analysis

A.4.1 Cross-language Generalization

ROIs	Spanish	Basque
IPL	0.245; moderate	0.816; anecdotal
LTL	0.452; anecdotal	0.461; anecdotal
VTL	5.37; moderate support of H1	6.277; moderate support of H1
dmPFC	0.316; anecdotal	0.359; anecdotal
IFG	0.241; moderate	2.72; anecdotal support of H1
vmPFC	0.249; moderate	1.061; anecdotal
PCG	0.221; moderate	0.255; moderate
ATL	0.745; anecdotal	0.197; moderate

Table A.5: Results of Bayesian analyses testing the evidence favor the null hypothesis in the cross-language generalization in the shallow condition, and the corresponding interpretation based on Lee and Wagenmakerś classification scheme. Regions in which the test moderately supported the alternative hypothesis (H1) are noted [3].

ROIs	Spanish	Basque
IPL	463.9; extreme	1902; extreme
LTL	124.1; extreme	8427; extreme
VTL	267.4; extreme	2165; extreme
dmPFC	40.91; very strong	66.76; very strong
IFG	1112; extreme	773.2; extreme
vmPFC	0.264; moderately support the null	5.961; moderate
PCG	4.389; moderate	108.8; extreme
ATL	0.214; moderately supports the null	0.194; moderately supports the null

Table A.6: Results of Bayesian analyses testing the evidence favor the alternative hypothesis in the cross-language generalization in the deep condition, and the corresponding interpretation based on Lee and Wagenmakerś classification scheme [3].

A.5 Cross-language Generalization with 15 ROIs

A set of 15 left-lateralized ROIs was pre-specified (see Figure A.4) based on a meta-analysis of the semantic system by Binder et al. 2009 [1] and one anterior temporal lobe (ATL) due to its crucial role as a "semantic hub" [15, 37, 72]. So, the ROIs included: inferior parietal lobe (IPL), inferior temporal lobe (ITL), middle temporal lobe (MTL), precuneus, fusiform gyrus (FFG), parahippocampal gyrus (PHG), superior frontal gyrus (SPG), posterior cingulate gyrus (PCG), pars opercularis (POP), pars triangularis (PTR), pars orbitalis (POR), frontal pole (FP), medial orbitofrontal cortex (MOFC), laterial orbitofrontal cortex (LOFC), and anterior temporal lobe (ATL).

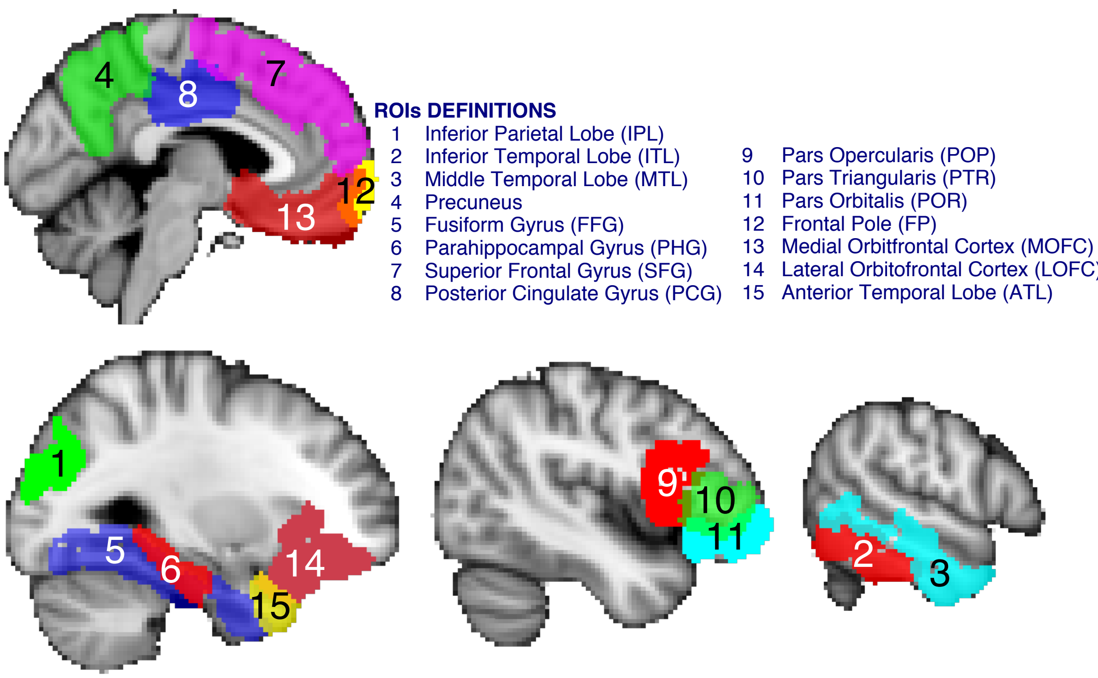


Figure A.4: The figure shows the selected regions of interest projected on an MNI standard template image. The 15 left-lateralized areas were pre-specified and included regions: inferior parietal lobe, inferior temporal lobe, middle temporal lobe, precuneus, fusiform gyrus, parahippocampal gyrus, superior frontal gyrus, posterior cingulate gyrus, pars opercularis, pars triangularis, pars orbitalis, frontal pole, medial orbitofrontal cortex, laterial orbitofrontal cortex and anterior temporal lobe.

It can be seen that in the shallow processing condition, the cross-language generalization from Basque to Spanish (see Figure A.6) was found to be not different from chance (FDR corrected for multiple comparisons) in all pre-specified ROIs including FP ($51.77 \pm 3.77; t(30) = 2.53; p = 0.26$), FFG ($51.71 \pm 4.41; t(30) = 2.09; p = 0.27$), IPL ($50.62 \pm 4.74; t(30) = 0.71 p = 0.74$), ITL ($50.76 \pm 4.14; t(30) = 0.99; p = 0.71$), LOFC ($50.39 \pm 3.81; t(30) = 0.55; p = 0.74$), MOFC ($50.36 \pm 4.70; t(30) = 0.42; p = 0.78$), MTL ($51.24 \pm 4.40; t(30) = 1.51; p = 0.42$), POP ($49.85 \pm 3.70; t(30) = -0.21; p = 0.86$), POR ($50.42 \pm 4.08; t(30) = 0.55; p = 0.74$), PTR ($50.47 \pm 3.88; t(30) = 0.65; p = 0.74$), PHG ($51.65 \pm 4.42; t(30) = 2.01; p = 0.27$), PCG ($49.84 \pm 4.76; t(30) = -0.18; p = 0.86$), Precuneus ($50.74 \pm 4.89; t(30) = 0.81; p = 0.74$), SFG ($50.86 \pm 4.43; t(30) = 1.05; p = 0.71$), ATL ($51.20 \pm 3.72; t(30) = 1.74; p = 0.34$). Similarly, the cross-language generalization from Spanish to Basque (see Figure A.5) was also found to be not different from chance (FDR corrected for multiple comparisons) in all pre-specified ROIs including FP ($51.53 \pm 4.33; t(30) = 1.91; p = 0.35$), FFG ($51.53 \pm 4.35; t(30) = 1.90; p = 0.35$), IPL ($51.66 \pm 4.97; t(30) = 1.80; p = 0.35$), ITL ($50.85 \pm 4.04; t(30) = 1.14; p = 0.38$), LOFC ($51.00 \pm 3.99; t(30) = 1.35; p = 0.38$), MOFC ($51.21 \pm 4.92; t(30) = 1.33; p = 0.38$), MTL ($51.16 \pm 3.87; t(30) = 1.61; p = 0.35$), POP ($51.31 \pm 4.173537473061428; t(30) = 1.69; p = 0.35$), POR ($50.40 \pm 4.35; t(30) = 0.49; p = 0.72$), PTR ($50.88 \pm 4.32; t(30) = 1.10; p = 0.38$), PHG ($50.93 \pm 4.04; t(30) = 1.24; p = 0.38$), PCG ($50.10 \pm 3.98; t(30) = 0.11; p = 0.91$), Precuneus ($50.72 \pm 4.97; t(30) = 0.78; p = 0.55$), SFG ($50.87 \pm 4.01; t(30) = 1.17; p = 0.38$), ATL ($50.12 \pm 3.99; t(30) = 0.16; p = 0.91$).

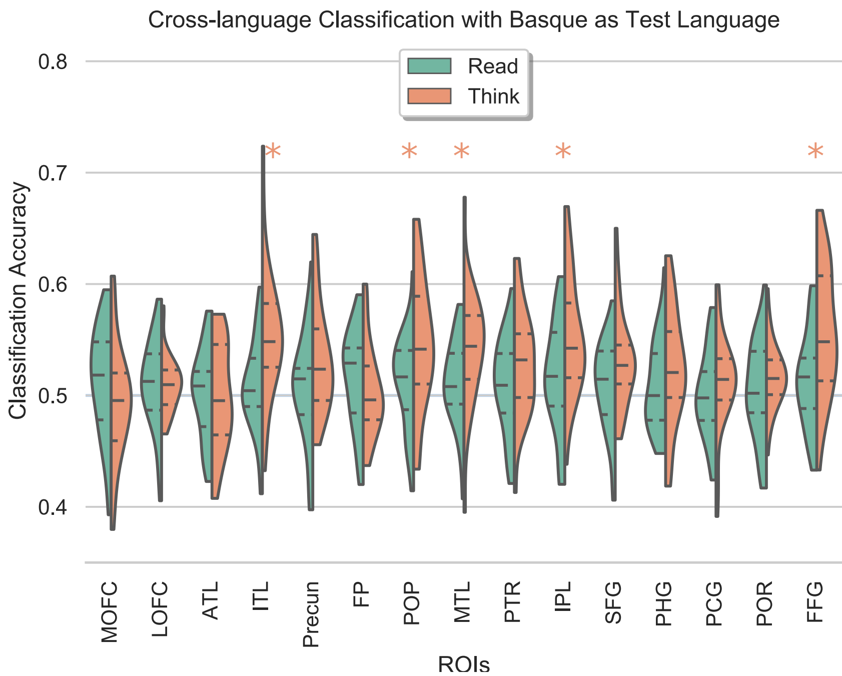


Figure A.5: The figure shows summary statistics of the ROIs for cross-language generalization from Spanish to Basque in both shallow and deep processing conditions. It can be seen that while the generalization was at chance-level in all ROIs in the shallow condition, it was statistically significantly above-chance and better than shallow in deep condition in five out of fifteen ROIs including FFG, IPL, MTL, POP and ITL. The three dotted lines inside each violin are the quartiles. The orange asterisks mark ROIs where cross-language generalization in deep was found to be statistically significantly above chance and better than shallow condition. The p-values were corrected for multiple comparisons.

In the deep processing condition on the other hand, the Basque to Spanish generalization (see Figure A.6) was found to be statistically significantly above-chance and better than shallow condition (FDR corrected for multiple comparisons) in 5 out of 15 ROIs including FP ($50.30 \pm 4.21; t(30) = 0.38; p = 0.76$), FFG ($56.19 \pm 6.50; t(30) = 5.12; p = 0.0003$), IPL ($54.50 \pm 5.12; t(30) = 4.74; p = 0.0004$), ITL ($54.48 \pm 5.57; t(30) = 4.33; p = 0.0006$), LOFC ($50.24 \pm 2.97; t(30) = 0.43; p = 0.76$), MOFC ($50.13 \pm 4.46; t(30) = 0.15; p = 0.88$), MTL ($53.50 \pm 5.52; t(30) = 3.41; p = 0.004$), POP ($54.53 \pm 5.52; t(30) = 4.41; p = 0.0006$), POR ($51.91 \pm 4.12; t(30) = 2.49; p = 0.03$), PTR ($53.10 \pm 3.98; t(30) = 4.19; p = 0.0007$), PHG ($51.97 \pm 5.46; t(30) = 1.95; p = 0.09$), PCG ($50.46 \pm 4.55; t(30) = 0.55; p = 0.76$), Precuneus $(52.98 \pm 5.56; t(30) = 2.89; p = 0.01)$, SFG $(53.04 \pm 4.39; t(30) = 3.73; p = 0.002)$, ATL $(49.64 \pm 4.28; t(30) = -0.45; p = 0.76)$. Similarly, Spanish to Basque generalization (see Figure A.5) was found to be statistically significantly above chance and better compared to shallow condition (FDR corrected for multiple comparisons) in five out of fifteen ROIs including: FP $(50.36 \pm 4.07; t(30) = 0.48; p = 0.68)$, FFG $(55.63 \pm 5.64; t(30) = 5.38; p = 8.41e - 05)$, IPL $(55.18 \pm 5.27; t(30) = 5.29; p = 8.41e - 05)$, ITL $(55.27 \pm 5.69; t(30) = 4.98; p = 0.0001)$, LOFC $(51.05 \pm 2.56; t(30) = 2.21; p = 0.048)$, MOFC $(49.45 \pm 5.04; t(30) = -0.59; p = 0.65)$, MTL $(54.23 \pm 5.06; t(30) = 4.50; p = 0.0004)$, POP $(54.43 \pm 5.86; t(30) = 4.08; p = 0.001)$, POR $(51.72 \pm 2.97; t(30) = 3.11; p = 0.007)$, PTR $(52.85 \pm 4.62; t(30) = 3.32; p = 0.005)$, PHG $(52.43 \pm 5.28; t(30) = 2.48; p = 0.030)$, PCG $(51.12 \pm 4.26; t(30) = 1.42; p = 0.21)$, Precuneus $(53.26 \pm 4.79; t(30) = 3.66; p = 0.002)$, SFG $(53.13 \pm 4.25; t(30) = 3.97; p = 0.001)$, ATL $(50.01 \pm 4.69; t(30) = 0.02; p = 0.99)$.

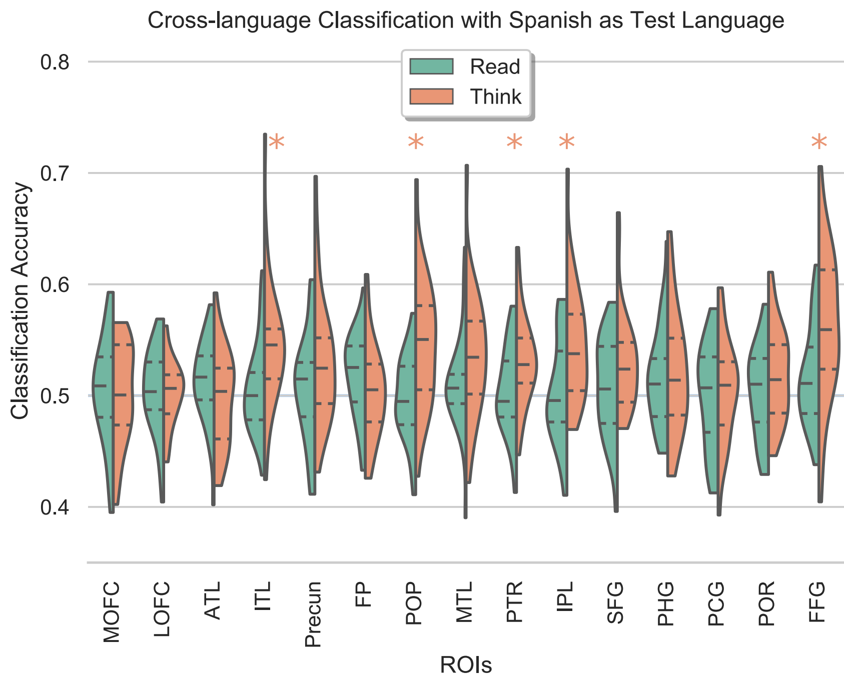


Figure A.6: The figure shows summary statistics of the ROIs for cross-language generalization from Basque to Spanish in both shallow and deep processing conditions. It can be seen that while the generalization was at chance-level in all ROIs in the shallow condition, it was statistically significantly above-chance and better than shallow in deep condition in five out of fifteen ROIs including FFG, IPL, PTR, POP and ITL. The three dotted lines inside each violin are the quartiles. The orange asterisks mark ROIs where cross-language generalization in deep was found to be statistically significantly above chance and better than shallow condition. The p-values were corrected for multiple comparisons.

A.6 Correlation between Cross-language Generalization and Language Proficiency

There were a few negative correlations between proficiency in Basque and Spanish indexed by the LeXTALE and cross-language generalization in LTL, IFG and dmPFC. However, these results should be taken with caution given that our study was not designed to explore inter-individual differences and that, while there were clear negative correlations, their statistical significance did not survive correction for multiple comparisons.

ROI	BEST scores	LeXTALE scores
IPL	$0.136; p = 0.497$	$-0.227; p = 0.255$
LTL	$-0.139; p = 0.489$	$-0.392; p = 0.043$
VTL	$-0.062; p = 0.759$	$-0.306; p = 0.120$
dmPFC	$-0.304; p = 0.123$	$-0.401; p = 0.038$
IFG	$-0.276; p = 0.164$	$-0.406; p = 0.036$
vmPFC	$-0.034; p = 0.866$	$-0.142; p = 0.479$
PCG	$-0.129; p = 0.522$	$-0.378; p = 0.052$
ATL	$0.026; p = 0.897$	$-0.282; p = 0.155$

Table A.7: The table shows correlation between cross-language generalization score, and the difference between proficiency scores between Basque and Spanish in the shallow condition. The p-values are uncorrected.

A.7 Semantic Analysis of the Stimuli

A matrix of word embeddings (word2vec) summarizing the semantic relationships between words within and across categories is presented in the Figure A.7.

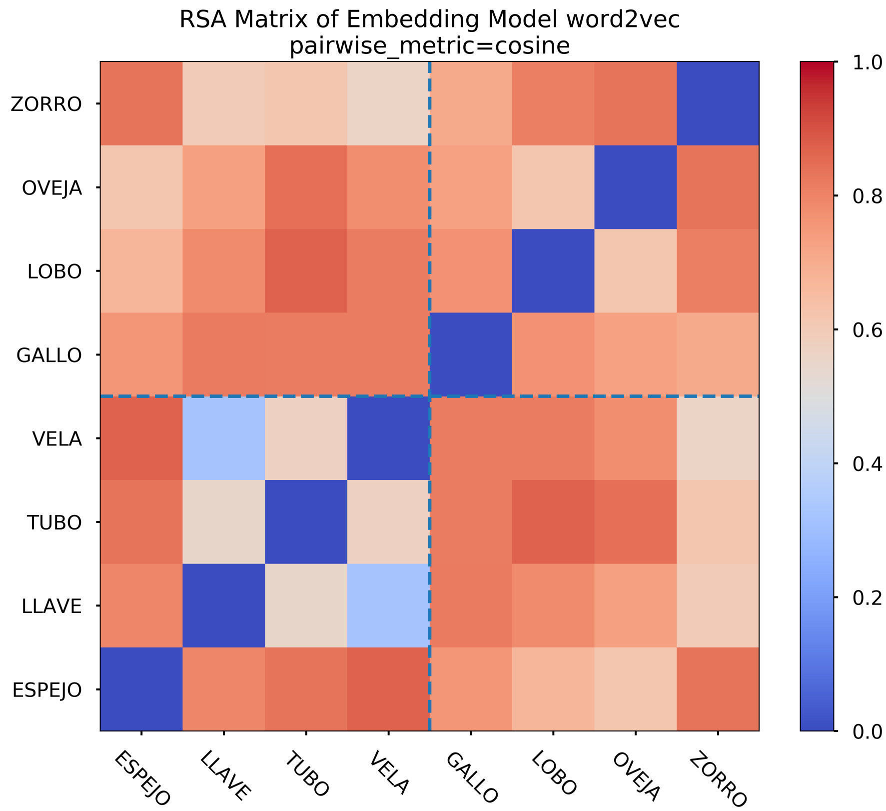


Figure A.7: The figure shows a matrix of word embeddings i.e. word2vec summarizing the semantic relationships between stimuli within and across categories.

Appendix B

Supplemental Materials for Chapter 4

B.1 List of Stimuli

#	Spanish	Basque	English	Category	Frequency	Concreteness
1	portero	atezain	doorman	animal	high	high
2	pintor	margolari	painter	animal	high	high
3	asesino	hiltzaile	murderer	animal	high	high
4	obispo	apezpiku	bishop	animal	high	high
5	gigante	erraldoi	giant	animal	high	high
6	periodista	kazetari	journalist	animal	high	high
7	lector	irakurle	reader	animal	high	high
8	secretario	idazkari	secretary	animal	high	high
9	abuelo	aitona	grandfather	animal	high	high
10	conductor	gidari	driver	animal	high	high
11	pescado	arrain	fish	animal	high	high
12	árbol	zuhaitz	tree	animal	high	high
13	criatura	izaki	creature	animal	high	high
14	muchacho	mutil	kid	animal	high	high
15	profesor	irakasle	teacher	animal	high	high
16	mujer	emakume	woman	animal	high	high

#	Spanish	Basque	English	Category	Frequency	Concreteness
17	corredor	lasterkari	runner	animal	high	low
18	rival	lehiakide	rival	animal	high	low
19	candidato	hautagai	candidate	animal	high	low
20	cliente	bezero	customer	animal	high	low
21	poeta	olerkari	poet	animal	high	low
22	edición	argitalpen	adviser	animal	high	low
23	hermana	arreba	sister	animal	high	low
24	jugador	jokalari	player	animal	high	low
25	presidente	lehendakari	president	animal	high	low
26	autor	egile	author	animal	high	low
27	experto	aditu	expert	animal	high	low
28	enemigo	etsai	enemy	animal	high	low
29	delegado	ordezkari	delegate	animal	high	low
30	testigo	lekuko	witness	animal	high	low
31	escritor	idazle	writer	animal	high	low
32	padre	aita	father	animal	high	low
33	fumador	erretzaile	smoker	animal	low	high
34	jinete	zaldizko	rider	animal	low	high
35	enfermera	erizain	nurse	animal	low	high
36	mensajero	mezulari	messenger	animal	low	high
37	caracol	barraskilo	snail	animal	low	high
38	yegua	behor	mare	animal	low	high
39	cocinero	sukaldari	cook	animal	low	high
40	cordero	arkume	lamb	animal	low	high
41	pandilla	koadrila	gang	animal	low	high
42	gallo	oilar	rooster	animal	low	high
43	peregrino	erromes	pilgrim	animal	low	high
44	cerdo	txerri	pig	animal	low	high
45	mariposa	tximeleta	butterfly	animal	low	high

#	Spanish	Basque	English	Category	Frequency	Concreteness
46	guardián	zaindari	guard	animal	low	high
47	roble	haritz	oak	animal	low	high
48	serpiente	suge	snake	animal	low	high
49	aprendiz	ikastun	apprentice	animal	low	low
50	idiota	memelo	idiot	animal	low	low
51	emisor	igorle	issuer	animal	low	low
52	inventor	asmatzaile	inventor	animal	low	low
53	pesimista	ezkor	pessimist	animal	low	low
54	confesor	aitorle	confessor	animal	low	low
55	portador	eramaile	bearer	animal	low	low
56	adivina	igarle	fortune-teller	animal	low	low
57	cabrón	aker	bastard	animal	low	low
58	usuario	erabiltzaile	user	animal	low	low
59	vendedor	saltzaile	salesman	animal	low	low
60	receptor	hartzaile	recipient	animal	low	low
61	ladrón	lapur	thief	animal	low	low
62	vencedor	garaile	victor	animal	low	low
63	invitado	gonbidatu	guest	animal	low	low
64	bruja	sorgin	witch	animal	low	low
65	sombra	itzal	shade	non-animal	high	high
66	página	orrialde	page	non-animal	high	high
67	regla	arau	rule	non-animal	high	high
68	juzgado	epaitegi	court	non-animal	high	high
69	taller	lantegi	workshop	non-animal	high	high
70	carretera	errepide	road	non-animal	high	high
71	viento	haize	wind	non-animal	high	high
72	anuncio	iragarpen	announcement	non-animal	high	high
73	restaurante	jatetxe	restaurant	non-animal	high	high
74	cerveza	garagardo	beer	non-animal	high	high
75	invierno	negu	winter	non-animal	high	high

#	Spanish	Basque	English	Category	Frequency	Concreteness
76	palacio	jauregi	palace	non-animal	high	high
77	moneda	txanpon	coin	non-animal	high	high
78	piedra	harri	stone	non-animal	high	high
79	nube	hodei	cloud	non-animal	high	high
80	firma	sinadura	signature	non-animal	high	high
81	contexto	testuinguru	context	non-animal	high	low
82	apertura	irekitze	opening	non-animal	high	low
83	fantasía	amets	fantasy	non-animal	high	low
84	olor	usain	smell	non-animal	high	low
85	comisión	batzorde	commission	non-animal	high	low
86	época	garai	time	non-animal	high	low
87	dicha	zorion	happiness	non-animal	high	low
88	accidente	istripu	accident	non-animal	high	low
89	posesión	edukitze	possession	non-animal	high	low
90	recurso	baliabide	resources	non-animal	high	low
91	asamblea	batzar	meeting	non-animal	high	low
92	trabajo	lan	job	non-animal	high	low
93	ahorro	aurrezki	saving	non-animal	high	low
94	triunfo	garaipen	victory	non-animal	high	low
95	lista	zerrenda	list	non-animal	high	low
96	ejercicio	ariketa	exercise	non-animal	high	low
97	estuche	kutxatila	case	non-animal	low	high
98	arroyo	erreka	stream	non-animal	low	high
99	cenicero	hautsontzi	ashtray	non-animal	low	high
100	posada	ostatu	sheet	non-animal	low	high
101	ceniza	errauts	ash	non-animal	low	high
102	ajedrez	xake	chess	non-animal	low	high
103	cinturón	gerriko	belt	non-animal	low	high
104	juguete	jostailu	toy	non-animal	low	high
105	concha	maskor	shell	non-animal	low	high
106	parcela	lursail	plot	non-animal	low	high

#	Spanish	Basque	English	Category	Frequency	Concreteness
107	fibra	zuntz	fiber	non-animal	low	high
108	guante	eskularru	glove	non-animal	low	high
109	espuma	apar	foam	non-animal	low	high
110	cortina	errezel	curtain	non-animal	low	high
111	anillo	eraztun	ring	non-animal	low	high
112	partícula	zatiki	particle	non-animal	low	high
113	vanidad	harrokeria	vanity	non-animal	low	low
114	ubicación	kokapen	location	non-animal	low	low
115	fascinación	lilura	fascination	non-animal	low	low
116	invención	asmaketa	invention	non-animal	low	low
117	robo	lapurreta	robbery	non-animal	low	low
118	curación	sendatze	treatment	non-animal	low	low
119	zumbido	burrunba	buzzing	non-animal	low	low
120	reglamento	araudi	rules	non-animal	low	low
121	censo	errolda	census	non-animal	low	low
122	cobijo	aterpe	shelter	non-animal	low	low
123	recado	mandatu	errand	non-animal	low	low
124	caucho	kautxu	rubber	non-animal	low	low
125	logro	lorpen	achievement	non-animal	low	low
126	olfato	usaimen	smell	non-animal	low	low
127	sequía	lehorte	drought	non-animal	low	low
128	engaño	iruzur	deception	non-animal	low	low

B.2 Detailed Cross-language Generation Results

#	ROI	CONDITION	MEAN	SD	t	p	$p*$
0	frontal pole	HFHC	59.27	4.34	8.802	0.0	0.0
1	frontal pole	HFLC	47.67	4.7	-2.04	0.057	3.429
2	frontal pole	LFHC	53.36	4.89	2.83	0.012	0.693
3	frontal pole	LFLC	45.14	4.8	-4.168	0.001	0.039
4	fusiform gyrus	HFHC	53.0	3.75	3.299	0.004	0.254
5	fusiform gyrus	HFLC	44.06	5.27	-4.642	0.0	0.014
6	fusiform gyrus	LFHC	54.96	5.33	3.837	0.001	0.079
7	fusiform gyrus	LFLC	45.16	4.55	-4.388	0.0	0.024
8	inferior parietal lobe	HFHC	52.79	4.16	2.767	0.013	0.791
9	inferior parietal lobe	HFLC	43.18	5.62	-5.005	0.0	0.007
10	inferior parietal lobe	LFHC	54.3	4.8	3.693	0.002	0.108
11	inferior parietal lobe	LFLC	44.58	3.22	-6.94	0.0	0.0
12	inferior temporal lobe	HFHC	52.39	3.77	2.612	0.018	1.093
13	inferior temporal lobe	HFLC	43.42	2.97	-9.14	0.0	0.0
14	inferior temporal lobe	LFHC	53.49	3.83	3.757	0.002	0.094
15	inferior temporal lobe	LFLC	46.11	4.01	-3.999	0.001	0.056
16	lateral orbitofrontal cortex	HFHC	54.72	3.48	5.602	0.0	0.002
17	lateral orbitofrontal cortex	HFLC	45.63	3.6	-5.007	0.0	0.006
18	lateral orbitofrontal cortex	LFHC	56.03	4.04	6.163	0.0	0.001
19	lateral orbitofrontal cortex	LFLC	44.97	4.47	-4.641	0.0	0.014
20	medial orbitofrontal cortex	HFHC	55.33	4.13	5.315	0.0	0.003
21	medial orbitofrontal cortex	HFLC	45.38	5.35	-3.56	0.002	0.144
22	medial orbitofrontal cortex	LFHC	53.82	2.76	5.705	0.0	0.002
23	medial orbitofrontal cortex	LFLC	43.62	5.11	-5.14	0.0	0.005
24	middle temporal lobe	HFHC	52.66	3.87	2.831	0.012	0.691
25	middle temporal lobe	HFLC	42.74	3.1	-9.645	0.0	0.0
26	middle temporal lobe	LFHC	53.14	3.15	4.118	0.001	0.043

27	middle temporal lobe	LFLC	46.35	3.1	-4.852	0.0	0.009
28	parahippocampal gyrus	HFHC	55.74	4.3	5.499	0.0	0.002
29	parahippocampal gyrus	HFLC	45.27	5.91	-3.3	0.004	0.254
30	parahippocampal gyrus	LFHC	54.16	4.87	3.523	0.003	0.157
31	parahippocampal gyrus	LFLC	44.4	4.94	-4.673	0.0	0.013
32	pars opercularis	HFHC	54.0	4.51	3.66	0.002	0.116
33	pars opercularis	HFLC	42.68	3.06	-9.876	0.0	0.0
34	pars opercularis	LFHC	53.64	6.03	2.489	0.023	1.407
35	pars opercularis	LFLC	46.56	4.03	-3.525	0.003	0.156
36	pars orbitalis	HFHC	56.77	4.34	6.436	0.0	0.0
37	pars orbitalis	HFLC	43.9	3.65	-6.884	0.0	0.0
38	pars orbitalis	LFHC	55.25	5.31	4.074	0.001	0.047
39	pars orbitalis	LFLC	46.13	5.26	-3.033	0.008	0.451
40	pars triangularis	HFHC	55.28	4.75	4.581	0.0	0.016
41	pars triangularis	HFLC	43.48	3.2	-8.395	0.0	0.0
42	pars triangularis	LFHC	55.48	5.2	4.342	0.0	0.027
43	pars triangularis	LFLC	46.47	5.08	-2.865	0.011	0.644
44	posterior cingulate gyrus	HFHC	57.15	5.04	5.846	0.0	0.001
45	posterior cingulate gyrus	HFLC	46.77	5.47	-2.434	0.026	1.574
46	posterior cingulate gyrus	LFHC	50.72	5.28	0.564	0.58	34.796
47	posterior cingulate gyrus	LFLC	45.35	4.74	-4.042	0.001	0.051
48	precuneus	HFHC	56.15	5.17	4.906	0.0	0.008
49	precuneus	HFLC	45.76	4.83	-3.622	0.002	0.126
50	precuneus	LFHC	51.25	4.27	1.203	0.246	14.736
51	precuneus	LFLC	44.24	4.37	-5.434	0.0	0.003
52	superior frontal lobe	HFHC	53.63	3.7	4.046	0.001	0.05
53	superior frontal lobe	HFLC	45.33	4.04	-4.766	0.0	0.011
54	superior frontal lobe	LFHC	54.42	3.12	5.836	0.0	0.001
55	superior frontal lobe	LFLC	44.44	4.92	-4.655	0.0	0.014
56	temporal pole	HFHC	57.06	4.17	6.985	0.0	0.0
57	temporal pole	HFLC	46.71	4.11	-3.296	0.004	0.256
58	temporal pole	LFHC	56.02	5.2	4.77	0.0	0.011
59	temporal pole	LFLC	46.6	5.38	-2.607	0.018	1.104

Bibliography

[1] J. R. Binder, R. H. Desai, W. W. Graves, and L. L. Conant, "Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies," *Cerebral Cortex*, vol. 19, no. 12, pp. 2767–2796, 2009.

[2] L. Chen, M. A. L. Ralph, and T. T. Rogers, "A unified model of human semantic knowledge and its disorders," *Nature human behaviour*, vol. 1, no. 3, p. 0039, 2017.

[3] M. D. Lee and E.-J. Wagenmakers, *Bayesian cognitive modeling: A practical course*. Cambridge university press, 2014.

[4] L. W. Barsalou, "Grounded cognition," *Annu. Rev. Psychol.*, vol. 59, pp. 617–645, 2008.

[5] J. R. Searle, "Minds, brains, and programs," *Behavioral and brain sciences*, vol. 3, no. 3, pp. 417–424, 1980.

[6] S. Harnad, "The symbol grounding problem," *Physica D: Nonlinear Phenomena*, vol. 42, no. 1-3, pp. 335–346, 1990.

[7] L. W. Barsalou, "Grounded cognition: Past, present, and future," *Topics in cognitive science*, vol. 2, no. 4, pp. 716–724, 2010.

[8] L. W. Barsalou, "On staying grounded and avoiding quixotic dead ends," *Psychonomic bulletin & review*, vol. 23, no. 4, pp. 1122–1142, 2016.

[9] G. Dove, "On the need for embodied and dis-embodied cognition," *Frontiers in Psychology*, vol. 1, p. 242, 2011.

[10] J. A. Fodor, *The language of thought*, vol. 5. Harvard university press, 1975.

[11] J. A. Fodor and J. A. Fodor, *Psychosemantics: The problem of meaning in the philosophy of mind*, vol. 2. MIT press, 1987.

[12] Z. W. Pylyshyn, "Computation and cognition: Issues in the foundations of cognitive science," *Behavioral and Brain Sciences*, vol. 3, no. 1, pp. 111–132, 1980.

[13] Z. W. Pylyshyn, "Computation and cognition," 1984.

[14] A. Newell, "Physical symbol systems," *Cognitive science*, vol. 4, no. 2, pp. 135–183, 1980.

[15] K. Patterson, P. J. Nestor, and T. T. Rogers, "Where do you know what you know? the representation of semantic knowledge in the human brain," *Nature Reviews Neuroscience*, vol. 8, no. 12, p. 976, 2007.

[16] R. L. Goldstone and L. W. Barsalou, "Reuniting perception and conception," *Cognition*, vol. 65, no. 2-3, pp. 231–262, 1998.

[17] E. K. Warrington and R. McCarthy, "Category specific access dysphasia," *Brain*, vol. 106, no. 4, pp. 859–878, 1983.

[18] E. K. Warrington and R. A. McCarthy, "Categories of knowledge: Further fractionations and an attempted integration," *Brain*, vol. 110, no. 5, pp. 1273–1296, 1987.

[19] E. K. Warrington and T. Shallice, "Category specific semantic impairments," *Brain*, vol. 107, no. 3, pp. 829–853, 1984.

[20] A. R. Damasio, "The brain binds entities and events by multiregional activation from convergence zones," *Neural computation*, vol. 1, no. 1, pp. 123–132, 1989.

[21] A. R. Damasio and H. Damasio, "Cortical systems for retrieval of concrete knowledge: The convergence zone framework.," 1994.

[22] L. W. Barsalou, "Perceptual symbol systems," *Behavioral and brain sciences*, vol. 22, no. 4, pp. 577–660, 1999.

[23] L. L. Chao, J. V. Haxby, and A. Martin, "Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects," *Nature neuroscience*, vol. 2, no. 10, pp. 913–919, 1999.

[24] L. L. Chao, J. Weisberg, and A. Martin, "Experience-dependent modulation of category-related cortical activity," *Cerebral Cortex*, vol. 12, no. 5, pp. 545–551, 2002.

[25] N. Kanwisher, J. McDermott, and M. M. Chun, "The fusiform face area: a module in human extrastriate cortex specialized for face perception," *Journal of neuroscience*, vol. 17, no. 11, pp. 4302–4311, 1997.

[26] J. J. Prinz, *Furnishing the mind: Concepts and their perceptual basis*. MIT press, 2004.

[27] J. R. Binder and R. H. Desai, "The neurobiology of semantic memory," *Trends in cognitive sciences*, vol. 15, no. 11, pp. 527–536, 2011.

[28] A. Martin, "The representation of object concepts in the brain," *Annu. Rev. Psychol.*, vol. 58, pp. 25–45, 2007.

[29] A. Martin and L. L. Chao, "Semantic memory and the brain: structure and processes," *Current opinion in neurobiology*, vol. 11, no. 2, pp. 194–201, 2001.

[30] W. K. Simmons, V. Ramjee, M. S. Beauchamp, K. McRae, A. Martin, and L. W. Barsalou, "A common neural substrate for perceiving and knowing about color," *Neuropsychologia*, vol. 45, no. 12, pp. 2802–2810, 2007.

[31] M. Kiefer, E.-J. Sim, B. Herrnberger, J. Grothe, and K. Hoenig, "The sound of concepts: Four markers for a link between auditory and conceptual brain systems," *Journal of Neuroscience*, vol. 28, no. 47, pp. 12224–12230, 2008.

[32] B. Z. Mahon and G. Hickok, "Arguments about the nature of concepts: Symbols, embodiment, and beyond," *Psychonomic bulletin & review*, vol. 23, no. 4, pp. 941–958, 2016.

[33] G. Dove, "Language as a disruptive technology: abstract concepts, embodiment and the flexible mind," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 373, no. 1752, p. 20170135, 2018.

[34] G. Dove, "Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition," *Psychonomic bulletin & review*, vol. 23, no. 4, pp. 1109–1121, 2016.

[35] G. Dove, "Beyond perceptual symbols: A call for representational pluralism," *Cognition*, vol. 110, no. 3, pp. 412–431, 2009.

[36] I. P. Kan, L. W. Barsalou, K. Olseth Solomon, J. K. Minor, and S. L. Thompson-Schill, "Role of mental imagery in a property verification task: fmri evidence for perceptual representations of conceptual knowledge," *Cognitive Neuropsychology*, vol. 20, no. 3-6, pp. 525–540, 2003.

[37] H. Damasio, T. J. Grabowski, D. Tranel, R. D. Hichwa, and A. R. Damasio, "A neural basis for lexical retrieval," *Nature*, vol. 380, no. 6574, p. 499, 1996.

[38] R. R. Davies, J. R. Hodges, J. J. Kril, K. Patterson, G. M. Halliday, and J. H. Xuereb, "The pathological basis of semantic dementia," *Brain*, vol. 128, no. 9, pp. 1984–1995, 2005.

[39] P. J. Nestor, T. D. Fryer, and J. R. Hodges, "Declarative memory impairments in alzheimer's disease and semantic dementia," *Neuroimage*, vol. 30, no. 3, pp. 1010–1020, 2006.

[40] B. Desgranges, V. Matuszewski, P. Piolino, G. Chételat, F. Mézenge, B. Landeau, V. De La Sayette, S. Belliard, and F. Eustache, "Anatomical and functional alterations in semantic dementia: a voxel-based mri and pet study," *Neurobiology of aging*, vol. 28, no. 12, pp. 1904–1913, 2007.

[41] W. K. Simmons and L. W. Barsalou, "The similarity-in-topography principle: Reconciling theories of conceptual deficits," *Cognitive neuropsychology*, vol. 20, no. 3-6, pp. 451–486, 2003.

[42] L. Fernandino, C. J. Humphries, M. S. Seidenberg, W. L. Gross, L. L. Conant, and J. R. Binder, "Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes," *Neuropsychologia*, vol. 76, pp. 17–26, 2015.

[43] J. Grainger, "Masked priming by translation equivalents in proficient bilinguals," *Language and cognitive processes*, vol. 13, no. 6, pp. 601–623, 1998.

[44] M. Perea, J. A. Dunabeitia, and M. Carreiras, "Masked associative/semantic priming effects across languages with highly proficient bilinguals," *Journal of Memory and Language*, vol. 58, no. 4, pp. 916–930, 2008.

[45] S. Schoonbaert, W. Duyck, M. Brysbaert, and R. J. Hartsuiker, "Semantic and translation priming from a first language to a second and back: Making sense of the findings," *Memory & cognition*, vol. 37, no. 5, pp. 569–586, 2009.

[46] J. F. Kroll and E. Stewart, "Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations," *Journal of memory and language*, vol. 33, no. 2, pp. 149–174, 1994.

[47] J. G. Van Hell and A. M. De Groot, "Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association," *Bilingualism: Language and cognition*, vol. 1, no. 3, pp. 193–211, 1998.

[48] J. Grainger and C. Beauvillain, "Associative priming in bilinguals: Some limits of interlingual facilitation effects.," *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 42, no. 3, p. 261, 1988.

[49] A. M. De Groot and G. L. Nas, "Lexical representation of cognates and noncognates in compound bilinguals," *Journal of memory and language*, vol. 30, no. 1, pp. 90–123, 1991.

[50] J. Altarriba and D. M. Basnight-Brown, "Methodological considerations in performing semantic-and translation-priming experiments across languages," *Behavior Research Methods*, vol. 39, no. 1, pp. 1–18, 2007.

[51] W. S. Francis, "Cognitive integration of language and memory in bilinguals: Semantic representation.," *Psychological bulletin*, vol. 125, no. 2, p. 193, 1999.

[52] Y.-S. Jin, "Effects of concreteness on cross-language priming in lexical decisions," *Perceptual and Motor Skills*, vol. 70, no. 3_suppl, pp. 1139–1154, 1990.

[53] A. Caramazza and I. Brones, "Semantic classification by bilinguals.," *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 34, no. 1, p. 77, 1980.

[54] R. Dufour and J. F. Kroll, "Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation," *Memory & Cognition*, vol. 23, no. 2, pp. 166–180, 1995.

[55] W. S. Francis and L. L. Goldmann, "Repetition priming within and between languages in semantic classification of concrete and abstract words," *Memory*, vol. 19, no. 6, pp. 653–663, 2011.

[56] T. Dijkstra and W. J. Van Heuven, "The architecture of the bilingual word recognition system: From identification to decision," *Bilingualism: Language and cognition*, vol. 5, no. 3, pp. 175–197, 2002.

[57] A. Mahmoudi, S. Takerkart, F. Regragui, D. Boussaoud, and A. Brovelli, "Multivoxel pattern analysis for fmri data: a review," *Computational and mathematical methods in medicine*, vol. 2012, 2012.

[58] "Language control in the bilingual brain," *Science*, vol. 312, no. 5779, pp. 1537–1540, 2006.

[59] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fmri data," *Trends in cognitive sciences*, vol. 10, no. 9, pp. 424–430, 2006.

[60] J.-D. Haynes, K. Sakai, G. Rees, S. Gilbert, C. Frith, and R. E. Passingham, "Reading hidden intentions in the human brain," *Current Biology*, vol. 17, no. 4, pp. 323–328, 2007.

[61] K. Jimura and R. A. Poldrack, "Analyses of regional-average activation and multivoxel pattern information tell complementary stories," *Neuropsychologia*, vol. 50, no. 4, pp. 544–552, 2012.

[62] R. A. Epstein and L. K. Morgan, "Neural responses to visual scenes reveals inconsistencies between fmri adaptation and multivoxel pattern analysis," *Neuropsychologia*, vol. 50, no. 4, pp. 530–543, 2012.

[63] T. Davis and R. A. Poldrack, "Measuring neural representations with fmri: practices and pitfalls," *Annals of the New York Academy of Sciences*, vol. 1296, no. 1, pp. 108–134, 2013.

[64] J. Illes, W. S. Francis, J. E. Desmond, J. D. Gabrieli, G. H. Glover, R. Poldrack, C. J. Lee, and A. D. Wagner, "Convergent cortical representation of semantic processing in bilinguals," *Brain and language*, vol. 70, no. 3, pp. 347–363, 1999.

[65] G. Ding, C. Perry, D. Peng, L. Ma, D. Li, S. Xu, Q. Luo, D. Xu, and J. Yang, "Neural mechanisms underlying semantic and orthographic processing in chinese–english bilinguals," *NeuroReport*, vol. 14, no. 12, pp. 1557–1562, 2003.

[66] J. J. Pillai, J. D. Allison, S. Sethuraman, J. M. Araque, D. Thiruvaiyaru, C. B. Ison, D. W. Loring, and T. Lavin, "Functional mr imaging study of language-related differences in bilingual cerebellar activation," *American journal of neuroradiology*, vol. 25, no. 4, pp. 523–532, 2004.

[67] L. A. Stowe and L. Sabourin, "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*, vol. 43, no. 4, pp. 329–353, 2005.

[68] P. Indefrey, "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*, vol. 56, pp. 279–304, 2006.

[69] M. W. Chee, N. Hon, H. L. Lee, and C. S. Soon, "Relative language proficiency modulates bold signal change when bilinguals perform semantic judgments," *Neuroimage*, vol. 13, no. 6, pp. 1155–1163, 2001.

[70] G. Xue, Q. Dong, Z. Jin, and C. Chen, "Mapping of verbal working memory in nonfluent chinese–english bilinguals with functional mri," *Neuroimage*, vol. 22, no. 1, pp. 1–10, 2004.

[71] A. Buchweitz, S. V. Shinkareva, R. A. Mason, T. M. Mitchell, and M. A. Just, "Identifying bilingual semantic neural representations across languages.," *Brain and language*, vol. 120, pp. 282–9, Mar 2012.

[72] J. a. Correia, E. Formisano, G. Valente, L. Hausfeld, B. Jansma, and M. Bonte, "Brain-based translation: fmri decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 34, pp. 332–8, Jan 2014.

[73] B. A. Shaywitz, S. E. Shaywitz, K. R. Pugh, W. E. Mencl, R. K. Fulbright, P. Skudlarski, R. T. Constable, K. E. Marchione, J. M. Fletcher, G. R. Lyon, *et al.*, "Disruption of posterior brain systems for reading in children with developmental dyslexia," *Biological psychiatry*, vol. 52, no. 2, pp. 101–110, 2002.

[74] B. McCandliss, L. Cohen, and S. Dehaene, "The visual word form area: expertise for reading in the fusiform gyrus," *Trends in cognitive sciences*, vol. 7, no. 7, pp. 293–299, 2003.

[75] S. Dehaene and L. Cohen, "The unique role of the visual word form area in reading," *Trends in cognitive sciences*, vol. 15, no. 6, pp. 254–262, 2011.

[76] J. R. Booth, D. D. Burman, F. W. V. Santen, Y. Harasaki, D. R. Gitelman, T. B. Parrish, and M. M. Mesulam, "The development of specialized brain systems in reading and oral-language," *Child Neuropsychology*, vol. 7, no. 3, pp. 119–141, 2001.

[77] G. Lerma-Usabiaga, M. Carreiras, and P. M. Paz-Alonso, "Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading," *Proceedings of the National Academy of Sciences*, p. 201803003, 2018.

[78] S. Dehaene and J.-P. Changeux, "Experimental and theoretical approaches to conscious processing," *Neuron*, vol. 70, no. 2, pp. 200–227, 2011.

[79] A. Y. Sklar, N. Levy, A. Goldstein, R. Mandel, A. Maril, and R. R. Hassin, "Reading and doing arithmetic nonconsciously," *Proceedings of the National Academy of Sciences*, vol. 109, no. 48, pp. 19614–19619, 2012.

[80] S. van Gaal and V. A. Lamme, "Unconscious high-level information processing implication for neurobiological theories of consciousness," *The neuroscientist*, vol. 18, no. 3, pp. 287–301, 2012.

[81] D. Soto and J. Silvanto, “Reappraising the relationship between working memory and conscious awareness,” *Trends in Cognitive Sciences*, vol. 18, no. 10, pp. 520–525, 2014.

[82] C. W. Eriksen, “Discrimination and learning without awareness: a methodological survey and evaluation.,” *Psychological review*, vol. 67, no. 5, p. 279, 1960.

[83] W. A. Johnston and V. J. Dark, “Selective attention,” *Annual review of psychology*, vol. 37, no. 1, pp. 43–75, 1986.

[84] A. G. Greenwald, “New look 3: Unconscious cognition reclaimed.,” *American Psychologist*, vol. 47, no. 6, p. 766, 1992.

[85] K. I. Forster, “The pros and cons of masked priming,” *Journal of psycholinguistic research*, vol. 27, no. 2, pp. 203–233, 1998.

[86] S. Kouider and E. Dupoux, “Partial awareness creates the ”illusion” of subliminal semantic priming,” *Psychological science*, vol. 15, no. 2, pp. 75–81, 2004.

[87] S. Kouider and S. Dehaene, “Levels of processing during non-conscious perception: a critical review of visual masking,” *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 362, no. 1481, pp. 857–875, 2007.

[88] D. G. Purcell, A. L. Stewart, and K. E. Stanovich, “Another look at semantic priming without awareness.,” *Perception & psychophysics*, vol. 34, pp. 65–71, Jul 1983.

[89] D. Holender, “Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal,” *Behavioral and brain Sciences*, vol. 9, no. 1, pp. 1–23, 1986.

[90] L. Naccache, R. Gaillard, C. Adam, D. Hasboun, S. Clémenceau, M. Baulac, S. Dehaene, and L. Cohen, “A direct intracranial record of emotions evoked by subliminal words.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 7713–7, May 2005.

[91] S. Dehaene, *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Penguin, 2014.

[92] S. Kouider, V. De Gardelle, J. Sackur, and E. Dupoux, “How rich is consciousness? the partial awareness hypothesis,” *Trends in cognitive sciences*, vol. 14, no. 7, pp. 301–307, 2010.

[93] A. J. Marcel, “Conscious and unconscious perception: Experiments on visual masking and word recognition,” *Cognitive psychology*, vol. 15, no. 2, pp. 197–237, 1983.

[94] J. T. Devlin, H. L. Jamison, P. M. Matthews, and L. M. Gonnerman, “Morphology and the internal structure of words,” *Proceedings of the National Academy of Sciences*, vol. 101, no. 41, pp. 14984–14988, 2004.

[95] R. L. Abrams and A. G. Greenwald, “Parts outweigh the whole (word) in unconscious analysis of meaning,” *Psychological science*, vol. 11, no. 2, pp. 118–124, 2000.

[96] R. L. Abrams, M. R. Klinger, and A. G. Greenwald, “Subliminal words activate semantic categories (not automated motor responses),” *Psychonomic Bulletin & Review*, vol. 9, no. 1, pp. 100–106, 2002.

[97] A. G. Greenwald, R. L. Abrams, L. Naccache, and S. Dehaene, “Long-term semantic memory versus contextual memory in unconscious number processing.,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 29, no. 2, p. 235, 2003.

[98] K. Nakamura, S. Dehaene, A. Jobert, D. Le Bihan, and S. Kouider, “Subliminal convergence of kanji and kana words: further evidence for functional parcellation of the posterior temporal cortex in visual word perception,” *Journal of Cognitive Neuroscience*, vol. 17, no. 6, pp. 954–968, 2005.

[99] N. F. Dixon, *Subliminal perception: The nature of a controversy*. McGraw-Hill, 1971.

[100] F. Schlaghecken, E. Blagrove, and E. A. Maylor, “No difference between conscious and nonconscious visuomotor control: evidence from perceptual learning in the masked prime task.,” *Consciousness and cognition*, vol. 17, pp. 84–93, Mar 2008.

[101] M. S. Pratte and J. N. Rouder, “A task-difficulty artifact in subliminal priming.,” *Attention, perception & psychophysics*, vol. 71, pp. 1276–83, Aug 2009.

[102] A. Lutz and E. Thompson, “Neurophenomenology integrating subjective experience and brain dynamics in the neuroscience of consciousness,” *Journal of consciousness studies*, vol. 10, no. 9-10, pp. 31–52, 2003.

[103] E. Van den Bussche, A. Vermeiren, K. Desender, W. Gevers, G. Hughes, T. Verguts, and B. Reynvoet, “Disentangling conscious and unconscious

processing: a subjective trial-based assessment approach.," *Frontiers in human neuroscience*, vol. 7, p. 769, 2013.

[104] S. J. Haase and G. D. Fisk, "Awareness of "invisible" arrows in a metacontrast masking paradigm," *The American journal of psychology*, vol. 128, no. 1, pp. 15–30, 2015.

[105] M. F. Damian, "Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 27, no. 1, p. 154, 2001.

[106] A. G. Greenwald, S. C. Draine, and R. L. Abrams, "Three cognitive markers of unconscious semantic activation," *Science*, vol. 273, no. 5282, pp. 1699–1702, 1996.

[107] S. C. Draine, *Analytic limitations of unconscious language processing*. PhD thesis, University of Washington, 1997.

[108] S. Dehaene, L. Naccache, G. Le Clec'H, E. Koechlin, M. Mueller, G. Dehaene-Lambertz, P.-F. van de Moortele, and D. Le Bihan, "Imaging unconscious semantic priming," *Nature*, vol. 395, no. 6702, p. 597, 1998.

[109] S. C. Draine and A. G. Greenwald, "Replicable unconscious semantic priming.," *Journal of Experimental Psychology: General*, vol. 127, no. 3, p. 286, 1998.

[110] L. Naccache and S. Dehaene, "Unconscious semantic priming extends to novel unseen stimuli.," *Cognition*, vol. 80, pp. 215–29, Jul 2001.

[111] R. Gaillard, A. Del Cul, L. Naccache, F. Vinckier, L. Cohen, and S. Dehaene, "Nonconscious semantic processing of emotional words modulates conscious access," *Proceedings of the National Academy of Sciences*, vol. 103, no. 19, pp. 7524–7529, 2006.

[112] A. Ohman and S. Mineka, "Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning," *Psychological Review*, vol. 108, pp. 483–522, Jul 2001.

[113] S. Mineka and A. Öhman, "Phobias and preparedness: The selective, automatic, and encapsulated nature of fear," *Biological Psychiatry*, vol. 52, no. 10, pp. 927–937, 2002.

[114] P. J. Whalen, J. Kagan, R. G. Cook, F. C. Davis, H. Kim, S. Polis, D. G. McLaren, L. H. Somerville, A. A. McLean, J. S. Maxwell, and T. Johnstone, "Human amygdala responsivity to masked fearful eye whites," *Science*, vol. 306, no. 5704, p. 2061, 2004.

[115] L. Carretié, J. A. Hinojosa, F. Mercado, and M. Tapia, "Cortical response to subjectively unconscious danger," *Neuroimage*, vol. 24, no. 3, pp. 615–623, 2005.

[116] V. Axelrod, M. Bar, G. Rees, and G. Yovel, "Neural correlates of subliminal language processing," *Cerebral Cortex*, vol. 25, no. 8, pp. 2160–2169, 2014.

[117] O. Hauk, I. Johnsrude, and F. Pulvermüller, "Somatotopic representation of action words in human motor and premotor cortex," *Neuron*, vol. 41, no. 2, pp. 301–307, 2004.

[118] K. Lemhöfer and M. Broersma, "Introducing lextale: A quick and valid lexical test for advanced learners of english," *Behavior research methods*, vol. 44, no. 2, pp. 325–343, 2012.

[119] A. De Bruin, M. Carreiras, and J. A. Duñabeitia, "The best dataset of language proficiency," *Frontiers in psychology*, vol. 8, p. 522, 2017.

[120] A. Duchon, M. Perea, N. Sebastián-Gallés, A. Martí, and M. Carreiras, "Espal: one-stop shopping for spanish word properties.," *Behavior research methods*, vol. 45, pp. 1246–58, Dec 2013.

[121] M. Perea, M. Urkia, C. J. Davis, A. Agirre, E. Laseka, and M. Carreiras, "E-hitz: a word frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (basque).," *Behavior research methods*, vol. 38, pp. 610–5, Nov 2006.

[122] S. V. Shinkareva, V. L. Malave, R. A. Mason, T. M. Mitchell, and M. A. Just, "Commonality of neural representations of words and pictures," *Neuroimage*, vol. 54, no. 3, pp. 2418–2425, 2011.

[123] J. W. Peirce, "Psychopy–psychophysics software in python.," *Journal of neuroscience methods*, vol. 162, pp. 8–13, May 2007.

[124] S. Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002.

[125] M. Jenkinson, P. Bannister, M. Brady, and S. M. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[126] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[127] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[128] M. Hanke, Y. O. Halchenko, P. B. Sederberg, S. J. Hanson, J. V. Haxby, and S. Pollmann, "Pymvpa: A python toolbox for multivariate pattern analysis of fmri data.," *Neuroinformatics*, vol. 7, no. 1, pp. 37–53, 2009.

[129] R. Fan, K. Chang, C. Hsieh, X. Wang, and C. Lin, "Liblinear: A library for large linear classification," *The Journal of Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.

[130] G. Varoquaux, P. R. Raamana, D. A. Engemann, A. Hoyos-Idrobo, Y. Schwartz, and B. Thirion, "Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines," *NeuroImage*, vol. 145, 06 2016.

[131] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fmri: a tutorial overview," *Neuroimage*, vol. 45, no. 1, pp. S199–S209, 2009.

[132] T. Mitchell, R. Hutchinson, R. Niculescu, F. Pereira, X. Wang, J. MA, and S. Newman, "Learning to decode cognitive states from brain images," *Mach. Learn.*, vol. 13, pp. 667–668, 01 2004.

[133] J. Mourão Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, "Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional mri data.," *NeuroImage*, vol. 28, pp. 980–95, Dec 2005.

[134] L. Weiskrantz, *Consciousness lost and found: A neuropsychological exploration.* Oxford University Press, 1997.

[135] A. Sahraie, L. Weiskrantz, J. Barbur, A. Simmons, S. Williams, and M. Brammer, "Pattern of neuronal activity associated with conscious and unconscious processing of visual signals," *Proceedings of the National Academy of Sciences*, vol. 94, no. 17, pp. 9406–9411, 1997.

[136] M. Overgaard, B. Timmermans, K. Sandberg, and A. Cleeremans, "Optimizing subjective measures of consciousness," *Consciousness and cognition*, vol. 19, no. 2, pp. 682–684, 2010.

[137] S. F. Cappa, D. Perani, T. Schnur, M. Tettamanti, and F. Fazio, "The effects of semantic category and knowledge type on lexical-semantic access: a pet study," *Neuroimage*, vol. 8, no. 4, pp. 350–359, 1998.

[138] M. Grossman, P. Koenig, C. DeVita, G. Glosser, D. Alsop, J. Detre, and J. Gee, "The neural basis for category-specific knowledge: an fmri study," *Neuroimage*, vol. 15, no. 4, pp. 936–948, 2002.

[139] J. Kounios, P. Koenig, G. Glosser, C. DeVita, K. Dennis, P. Moore, and M. Grossman, "Category-specific medial temporal lobe activation and the consolidation of semantic memory: evidence from fmri," *Cognitive Brain Research*, vol. 17, no. 2, pp. 484–494, 2003.

[140] T. Wheatley, J. Weisberg, M. S. Beauchamp, and A. Martin, "Automatic priming of semantically related words reduces activity in the fusiform gyrus.," *Journal of cognitive neuroscience*, vol. 17, pp. 1871–85, Dec 2005.

[141] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. A. Mason, and M. A. Just, "Predicting human brain activity associated with the meanings of nouns," *science*, vol. 320, no. 5880, pp. 1191–1195, 2008.

[142] M. A. Just, V. L. Cherkassky, S. Aryal, and T. M. Mitchell, "A neurosemantic theory of concrete noun representation based on the underlying brain codes," *PloS one*, vol. 5, no. 1, p. e8622, 2010.

[143] S. Dehaene, M. Kerszberg, and J.-P. Changeux, "A neuronal model of a global workspace in effortful cognitive tasks," *Proceedings of the National Academy of Sciences*, vol. 95, no. 24, pp. 14529–14534, 1998.

[144] S. Dehaene, L. Naccache, L. Cohen, D. Le Bihan, J.-F. Mangin, J.-B. Poline, and D. Rivière, "Cerebral mechanisms of word masking and unconscious repetition priming," *Nature neuroscience*, vol. 4, no. 7, p. 752, 2001.

[145] S. Kouider, S. Dehaene, A. Jobert, and D. Le Bihan, "Cerebral bases of subliminal and supraliminal priming during reading," *Cerebral Cortex*, vol. 17, no. 9, pp. 2019–2029, 2006.

[146] S. Van Gaal, L. Naccache, J. D. Meuwese, A. M. Van Loon, A. H. Leighton, L. Cohen, and S. Dehaene, "Can the meaning of multiple words be integrated unconsciously?," *Phil. Trans. R. Soc. B*, vol. 369, no. 1641, p. 20130212, 2014.

[147] K. Eo, O. Cha, S. C. Chong, and M.-S. Kang, "Less is more: semantic information survives interocular suppression when attention is diverted," *Journal of Neuroscience*, vol. 36, no. 20, pp. 5489–5497, 2016.

[148] T. Heyman and P. Moors, "Using interocular suppression and eeg to study semantic processing," *Journal of Neuroscience*, vol. 32, no. 5, pp. 1515–1516, 2012.

[149] M.-S. Kang, R. Blake, and G. F. Woodman, "Semantic analysis does not occur in the absence of awareness induced by interocular suppression," *Journal of Neuroscience*, vol. 31, no. 38, pp. 13535–13545, 2011.

[150] M. T. Diaz and G. McCarthy, "Unconscious word processing engages a distributed network of brain regions," *Journal of Cognitive Neuroscience*, vol. 19, no. 11, pp. 1768–1775, 2007.

[151] F. Bergström and J. Eriksson, "Maintenance of non-consciously presented information engages the prefrontal cortex," *Frontiers in human neuroscience*, vol. 8, p. 938, 2014.

[152] F. Bergström and J. Eriksson, "Neural evidence for non-conscious working memory," *Cerebral Cortex*, pp. 1–12, 2017.

[153] A. Dutta, K. Shah, J. Silvanto, and D. Soto, "Neural basis of non-conscious visual working memory," *Neuroimage*, vol. 91, pp. 336–343, 2014.

[154] B. Van Vugt, B. Dagnino, D. Vartak, H. Safaai, S. Panzeri, S. Dehaene, and P. R. Roelfsema, "The threshold for conscious report: Signal loss and response bias in visual and frontal cortex," *Science*, vol. 360, no. 6388, pp. 537–542, 2018.

[155] B. Zinszer, A. J. Anderson, O. Kang, T. Wheatley, and R. D. Raizada, "You say potato, i say tǔdòu: How speakers of different languages share the same concept.," in *CogSci*, 2015.

[156] M. Dehghani, R. Boghrati, K. Man, J. Hoover, S. I. Gimbel, A. Vaswani, J. D. Zevin, M. H. Immordino-Yang, A. S. Gordon, A. Damasio, *et al.*, "Decoding the neural representation of story meanings across languages," *Human brain mapping*, vol. 38, no. 12, pp. 6096–6106, 2017.

[157] J. N. Williams, "Is automatic priming semantic?," *European Journal of Cognitive Psychology*, vol. 8, no. 2, pp. 113–162, 1996.

[158] D. M. Basnight-Brown and J. Altarriba, "Differences in semantic and translation priming across languages: The role of language direction and language dominance," *Memory & cognition*, vol. 35, no. 5, pp. 953–965, 2007.

[159] D. A. Balota and J. I. Chumbley, "Are lexical decisions a good measure of lexical access? the role of word frequency in the neglected decision stage.," *Journal of Experimental Psychology: Human perception and performance*, vol. 10, no. 3, p. 340, 1984.

[160] J. Abutalebi, S. F. Cappa, and D. Perani, "The bilingual brain as revealed by functional neuroimaging," *Bilingualism: Language and cognition*, vol. 4, no. 2, pp. 179–190, 2001.

[161] M. W. Chee, C. S. Soon, and H. L. Lee, "Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation," *Journal of Cognitive Neuroscience*, vol. 15, no. 1, pp. 85–97, 2003.

[162] F. I. Craik and R. S. Lockhart, "Levels of processing: A framework for memory research," *Journal of Verbal Learning and Verbal Behavior*, vol. 11, pp. 671–684, Dec. 1972.

[163] R. H. Pruim, M. Mennes, D. van Rooij, A. Llera, J. K. Buitelaar, and C. F. Beckmann, "Ica-aroma: A robust ica-based strategy for removing motion artifacts from fmri data," *Neuroimage*, vol. 112, pp. 267–277, 2015.

[164] E.-J. Wagenmakers, J. Love, M. Marsman, T. Jamil, A. Ly, J. Verhagen, R. Selker, Q. F. Gronau, D. Dropmann, B. Boutin, *et al.*, "Bayesian inference for psychology. part ii: Example applications with jasp," *Psychonomic bulletin & review*, vol. 25, no. 1, pp. 58–76, 2018.

[165] J. Team *et al.*, "Jasp (version 0.8. 6)[computer software]," 2018.

[166] O. Yamashita, M.-a. Sato, T. Yoshioka, F. Tong, and Y. Kamitani, "Sparse estimation automatically selects voxels relevant for the decoding of fmri activity patterns," *NeuroImage*, vol. 42, no. 4, pp. 1414–1429, 2008.

[167] S. Van Schooten, R. Harel, S. Ercan, and E. De Groot, "Applying feature selection methods on fmri data," *Student project report*, 2014.

[168] D. Soto, U. A. Sheikh, N. Mei, and R. Santana, "Decoding and encoding models reveal the role of the depth of processing in the brain representation of meaning," *bioRxiv*, 2019.

[169] E. Van de Putte, W. De Baene, M. Brass, and W. Duyck, "Neural overlap of l1 and l2 semantic representations in speech: A decoding approach," *NeuroImage*, vol. 162, pp. 106–116, 2017.

[170] E. Van de Putte, *The representation of language in bilinguals: Neural overlap as a function of modality, representational level, language proficiency and context.* PhD thesis, Ghent University, 2018.

[171] J. A. Stolz and D. Besner, "On the myth of automatic semantic activation in reading," *Current Directions in Psychological Science*, vol. 8, pp. 61–65, Apr. 1999.

[172] M. A. L. Ralph, E. Jefferies, K. Patterson, and T. T. Rogers, "The neural and computational bases of semantic cognition," *Nature Reviews Neuroscience*, vol. 18, no. 1, p. 42, 2017.

[173] M. A. Lambon Ralph, "Neurocognitive insights on conceptual knowledge and its breakdown," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1634, p. 20120392, 2014.

[174] A. R. Luria, *The working brain: An introduction to neuropsychology.* USA: Basic Books, 1976.

[175] S. L. Thompson-Schill, M. D'Esposito, G. K. Aguirre, and M. J. Farah, "Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation," *Proceedings of the National Academy of Sciences*, vol. 94, no. 26, pp. 14792–14797, 1997.

[176] A. D. Wagner, E. J. Paré-Blagoev, J. Clark, and R. A. Poldrack, "Recovering meaning: left prefrontal cortex guides controlled semantic retrieval," *Neuron*, vol. 31, no. 2, pp. 329–338, 2001.

[177] K. A. Noonan, E. Jefferies, F. Corbett, and M. A. Lambon Ralph, "Elucidating the nature of deregulated semantic cognition in semantic aphasia: evidence for the roles of prefrontal and temporo-parietal cortices," *Journal of Cognitive Neuroscience*, vol. 22, no. 7, pp. 1597–1613, 2010.

[178] A. Rodriguez-Fornells, M. Rotte, H.-J. Heinze, T. Nösselt, and T. F. Münte, "Brain potential and functional mri evidence for how to handle two languages with one brain," *Nature*, vol. 415, no. 6875, p. 1026, 2002.

[179] L. J. Hoversten, T. Brothers, T. Y. Swaab, and M. J. Traxler, "Language membership identification precedes semantic access: Suppression during bilingual word recognition," *Journal of Cognitive Neuroscience*, vol. 27, no. 11, pp. 2108–2116, 2015.

[180] G. Oppenheim, Y. J. Wu, and G. Thierry, "Found in translation: Late bilinguals do automatically activate their native language when they are not using it," *Cognitive science*, vol. 42, no. 5, pp. 1700–1713, 2018.

[181] M. Brysbaert, I. Van Wijnendaele, and W. Duyck, "On the temporal delay assumption and the impact of non-linguistic context effects," *Bilingualism: Language and Cognition*, vol. 5, no. 3, pp. 199–201, 2002.

[182] L. Demetriou, O. S. Kowalczyk, G. Tyson, T. Bello, R. D. Newbould, and M. B. Wall, "A comprehensive evaluation of multiband-accelerated sequences

and their effects on statistical outcome measures in fmri," *BioRxiv*, p. 076307, 2016.

[183] L. Chen, A. T. Vu, J. Xu, S. Moeller, K. Ugurbil, E. Yacoub, and D. A. Feinberg, "Evaluation of highly accelerated simultaneous multi-slice epi for fmri," *Neuroimage*, vol. 104, pp. 452–459, 2015.

[184] H. Goodglass, M. R. Hyde, and S. Blumstein, "Frequency, picturability and availability of nouns in aphasia," *Cortex*, vol. 5, no. 2, pp. 104–119, 1969.

[185] E. Jefferies, S. S. Baker, M. Doran, and M. A. L. Ralph, "Refractory effects in stroke aphasia: a consequence of poor semantic control," *Neuropsychologia*, vol. 45, no. 5, pp. 1065–1079, 2007.

[186] A. M. de Groot, "Representational aspects of word imageability and word frequency as assessed through word association.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, no. 5, p. 824, 1989.

[187] C. T. James, "The role of semantic information in lexical decisions.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 1, no. 2, p. 130, 1975.

[188] A. Paivio, "Dual coding theory: Retrospect and current status.," *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 45, no. 3, p. 255, 1991.

[189] S. Monsell, M. C. Doyle, and P. N. Haggard, "Effects of frequency on visual word recognition tasks: Where are they?," *Journal of Experimental Psychology: General*, vol. 118, no. 1, p. 43, 1989.

[190] C. J. Davis and M. Perea, "Buscapalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in spanish," *Behavior Research Methods*, vol. 37, no. 4, pp. 665–671, 2005.

[191] J. Duñabeitia, A. Casaponsa, M. Dimitropoulou, A. Martí, S. Larraza, and M. Carreiras, "Basp: A basque–spanish database of translation equivalents," *Manuscript in preparation*, 2017.

[192] M. Coltheart, "The mrc psycholinguistic database," *The Quarterly Journal of Experimental Psychology Section A*, vol. 33, no. 4, pp. 497–505, 1981.

[193] N. Kriegeskorte, R. Goebel, and P. Bandettini, "Information-based functional brain mapping," *Proceedings of the National Academy of Sciences*, vol. 103, no. 10, pp. 3863–3868, 2006.