

ECOLE NORMALE SUPERIEURE
UNIVERSITE PARIS DESCARTES
ECOLE DES HAUTES ETUDES
EN SCIENCES SOCIALES

MASTER THESIS

**Similarity and Association:
Principles of Distributed Semantic
Processing in the Human Brain**

Author:
Songsheng YING

Supervisors:
Sabine PLOUX
Laurent BONNASSE-GAHOT

Co-supervisor:
Christophe PALLIER

*A thesis submitted in fulfillment of the requirements
for the degree of
Research Master in Cognitive Science*

*Research project executed in the
Centre for Social Analysis and Mathematics
CAMS, UMR 8557 CNRS-EHESS*

June 7, 2019

Declaration of Originality

This master's project consists of an original research on the dissociation of multiple organization principles of human semantic processing. Namely, we relate various language-related cortical areas with different semantic functions using an fMRI encoding experiments.

This study differs from existing works in computational linguistics in the following points.

- It attempts to build non-generic semantic word embeddings, targeting specific semantic principles parallel to the paradigmatic and syntagmatic axes. Thus it requires fine tuning certain parameters when configuring existing embedding generation algorithms.
- Mathematic operations are applied on different types of semantic spaces to extract new embedding spaces. These manipulations show interactions between different embeddings.
- New baseline benchmark datasets are made available for French word-pair semantic proximity evaluation.

This study further investigates current theories and discoveries on a hypothetical function locus of semantic processing and multiple cortices' contribution in verbal comprehension with fMRI data.

- Existing works in semantic fMRI encoding either use non-ecological stimuli to reveal semantic condition contrasts, or ecological stimuli to reveal general semantic processing without targeting different (hypothetical) semantic aspects.
- A possible construction of *similarity* and *association* semantic memories are proposed and tested, which is based on converging evidences and theories on paradigmatic axis/semantic hub/convergence zone and syntagmatic axis/ associational activation.
- The anterior temporal lobe localization hypothesis for a central semantic processing component is tested with ecological fMRI encoding.
- This is the first project analyzing French fMRI data collected in the project "Neural Computational Models of Natural Language" (PI: John Hale and Christophe Pallier). Precedent projects were performed with English data.

The fMRI encoding pipeline also differs from most other works in the following aspects.

- When training voxel-models for BOLD prediction, a large grid-search for best regression parameters is carried out so that each voxel is modeled by the most appropriate functional features.
- Multiple condition contrasting methods are employed including nested-model improvement testing and non-nested model performance comparison.

Declaration of Contribution

I, Songsheng YING(**SY**), declare that this Master's thesis titled, "Similarity and Association: Principles of Distributed Semantic Processing in the Human Brain" and the work presented in it could not have been accomplished without the help from the advisors of my research internship: Sabine PLOUX (**SP**), Laurent BONNASSE-GAHOT (**LBG**) and Christophe PALLIER (**CP**).

- **SP** helped enormously with the initial research problem definition and hypothesis formulation. Together with **LBG** and **CP**, they've given a rich literature collection on the relevant domain including brain encoding/decoding, theories on paradigmatic/syntagmatic axes.
- **CP** and Snezana TODOROVIC designed the MRI experiment, prepared audio stimuli, textual reference data and behavioral control procedures. **CP** preprocessed MRI data including data cleaning and normalization. These acquisitions were performed as part of the project "Neural Computational Models of Natural Language", financed by the NSF-ANR (PI: John Hale and Christophe Pallier).
- **SY** implemented automatic pipelines and realized a large part of rule-based and manual correction of transcribed fMRI stimuli lemmatisation and semantic space vocabulary alignment. **SP** helped the verification of the results.
- **SY** adopted multiple semantic model validation tasks and dataset, implemented the initial iteration of dataset translation and correction, which is perfected by **SP**.
- **SY** adapted publicly available algorithms and tested combinations of hyperparameter configurations to build one of the semantic representation models.
- **LBG** actively participated in the conception of semantic space dissociation algorithm, which is later implemented and tested by **SY**.
- **SP** visually examined iterations of the resulting semantic embedding models and assured their quality and coherence.
- **CP** shared the fMRI analysis code base developed by him and collaborators. This project heavily depended on the regressor generation, design matrix orthogonalization functions of the library. **CP** also shared the computed feature of acoustic energy. **SY** added upon the code base functions to perform nested GridSearch regressions, implemented multiple statistical tests and result visualization pipelines.
- **LBG** and **CP** actively participated in the analysis and interpretation of fMRI regression results. They also gave useful guidance on the utilization of `nilearn`, `nibabel`, `nistat` libraries and analytical procedures including statistical tests and ROI analysis.
- **SY** drafted the Master's thesis, which is attentively proofread by **SP**, who also gave thoughtful insights on psycholinguistic discussion of the obtained results. **LBG** offered countless redactional good-practice advice during the drafting, which helped shaping this thesis.

Songsheng YING

Supervisor: Sabine Ploux

Laboratory: Centre d'analyse et de mathématique sociales

Intended Memory Defense Session: July 2019

Language: English

Suggested Reviewer: [TODO]

Similarity and Association: Principles of Distributed Organisation of Semantics in the Human Brain

Introduction

Background. General semantic knowledge associates verbal and non-verbal stimuli to concepts internal to human cognitive system. In human language processing, lexicons with or without contexts are linked to their meanings by the lexicon semantic system. How brain processes semantics remains an open question. Tentatives to localize a stable semantic memory lead neuro- and computational linguists to a hub-and-spoke model (see Ralph, Jefferies, Patterson, & Rogers, 2017 for a review). A neural architecture of transmodal semantic memory across concepts with similar semantic significance, with the locus centered on bilateral ventrolateral anterior temporal lobe (vLATL) is suggested by pathological studies on *semantic dementia* (SD), *herpes simplex virus encephalitis* (HSVE) and other semantic disorders (Patterson, Nestor, & Rogers, 2007). While semantics' relevance to perception and action suggest a widely distributed, modality-specific neural network such as visual cortices (Borghesani et al., 2016). Pereira et al. (2018) built a BOLD-to-word decoder with GloVe (Pennington, Socher, & Manning, 2014), Huth, Nishimoto, Vu, & Gallant (2012) and Huth, De Heer, Griffiths, Theunissen, & Gallant (2016) used a 985-dimensional word-level co-occurrence based embedding space and narrative-story listening functional magnetic resonance imaging (fMRI) to build association maps. These results found an extensively distributed informative voxels in language, task, visual and other networks.

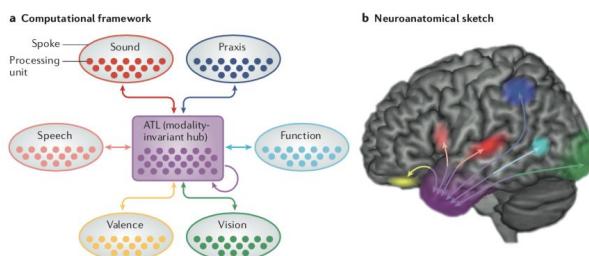


Figure 1. Adopted from (Ralph, Jefferies, Patterson, & Rogers, 2017), Hub-and-spoke model. a. Modality-specific informations are encoded separately in different processing layers. Such information is fed to a transmodal semantic hub, which contains conceptual knowledge, which reactivates complementary information in other spoke layers. b. A neuroanatomical representation of the

hub-and-spoke model, where the hub is located near ATL, spoke components distributed across different cortices.

(Peelen & Caramazza, 2012) used conceptual/perceptual contrast in a fMRI study to confirm that the conceptual information is most likely to be stored in BA20, BA38 (ATL), while perceptual information in posterior occipitotemporal cortex. Studies on HSVE compared with SD shows an intact performance at basic semantic levels (such as dog, knife) but not at subordinate level (for example poodle or bread knife), suggesting the ATL hub might be organizing concepts hierarchically.

Rationale. This project tends to build a semantic mapping based on hub-and-spoke model, on explicitly defining the feature space of hub and spoke-components.

Key research question. What are the semantic information encoded by the semantic hub and other components?

General hypotheses. Semantic hub encodes conceptual-similarity by conceptual super-subordinate hierarchies, while other components encode perceptual-specificities depending on their modality and all other associative relations. The semantic hub is localized near left and/or right ATL, while other components correspond to other cortical areas.

Methods

Model

Consistent with (Lofi, 2015), we define conceptual-similarity in terms of taxonomical properties (such as *cat/tiger*, *museum/theater*, but not *computer/software*, as computer is a type of hardware), and association in terms of relevance of two concepts (such as *computer/software*). In order to dissociate semantic hub activation patterns from other components, we propose a novel word-embedding scheme that rejects conceptual-similarity but keep association to build a semantic encoding model for fMRI data. This new embedding space, together with conceptual-similarity embedding space and classic statistical word-embeddings will be separately tested on similarity and association benchmarks if the dissociation of two aspects is indeed implemented.

Due to the lack of availability of benchmarks in French, we first build the embedding models in English to test the dissociation algorithm, then replicate the method with French data for fMRI encoding.

Todorovic & Pallier (2018) built a word-fMRI encoding model with GloVe and English-stimuli fMRI data. Similarly, we will construct regression-based machine learning models to encode word-embedding vectors into individual voxel BOLD signals of each participant with French fMRI. We will then identify and interpret the systematic differences of voxel activation profile.

Key features. We will be comparing the encoding performance for each voxel of three different embedding spaces. All three embeddings contain non-semantic dimensions including auditory signal existence, word-speed, acoustic signal energy and bottom-up syntax parser. The conceptual-similarity is word-embedding space constructed from WordNet, which is a tree-structure ontology organized by synonym sets and super-subordinate relations. The classic statistical word-embedding is adapted from GloVe. The pure association embedding is based on GloVe, but decorrelated with the conceptual-similarity space, which is presumed to encode only semantic associations.

Rationale of feature selection. The semantic hub is hypothesized to encode conceptual-similarity. While traditional word-embedding encoding and decoding studies found a distributed mapping between brain regions and semantic vectors, including the loci of semantic hub. We want to test if conceptual-information-based embeddings match better with semantic hub activations, and if traditional embeddings' encoding performance near the semantic hub region, would decrease significantly after decorrelation with conceptual-similarity embeddings.

Programming language. We will use Python 3 to build the decorrelation algorithm.

External scripts. As one candidate for semantic conceptual-similarity embeddings, WordNet embeddings will be constructed using algorithm provided by (Saedi, Branco, António Rodrigues, & Silva, 2018). We will build such space considering only synonymy, hypernymy, hyponymy, verb participle, adjective/adverb derivation and pertainym relationships in WordNet. Meronyms, holonyms and other relationships are rejected as they are more associative. The selected vocabulary will overlap at maximum with audio stimuli provided to fMRI recording participants.

The embedding to fMRI encoding regression algorithm is implemented by Verdier, Lakretz, & Pallier (2018).

Assumptions. We assume that conceptual-similarity information encoded by classic word-embeddings is contained by conceptual-similarity embeddings, such that after decorrelation process, the residual embedding space would comprise only non-conceptual (thus purely association) data.

Input data

Embedding construction. For conceptual-similarity embeddings, we will use English and French WordNet as source data to build WordNet embeddings (Miller, 1995), (Pradet, De Chalendar, & Baguenier-Desormeaux, 2014). They are thesaurus-like database organised hierarchically based on super-subordinate relations. In addition we will also test the performance of (Saedi, Branco, António Rodrigues, & Silva, 2018)'s algorithm with synonym databases, which are available in English and French, created by thesauri fusion and symmetrisation (Ploux & Ji, 2003).

For classic word-embeddings, we use GloVe embeddings that are provided with open access¹. They are co-occurrence frequency based statistical measures derived dense semantic vectorial representations.

The non-semantic data will be provided by Todorovic & Pallier (2018).

Embedding validation. With built and decorrelated word-embedding models, we will use vectorial distance to evaluate word-pair similarity and association with multiple benchmarking datasets (Lofi, 2015). For conceptual-similarity benchmarks, we use datasets provided by (Rubenstein & Goodenough, 1965), (Agirre et al., 2009) and (Hill, Reichart, & Korhonen, 2015). For association benchmarks, there is few available datasets due to the less clear definition of association (or relatedness in other terms), we adapt datasets from (Agirre et al., 2009) and (Halawi, Dror, Gabrilovich, & Koren, n.d.). The benchmarks are word pairs associated with a similarity or association score. Vectorial distance scores will be matched against benchmarks with pearson and spearman correlation.

¹ English GloVe: <https://nlp.stanford.edu/projects/glove/>, French DepGloVe with lemma: <http://alpage.inria.fr/frmgwiki/content/word-embeddings-avec-depglove>

fMRI data. We will be using fMRI data acquired in (Todorovic & Pallier, 2018), in which 20 native French speakers listen to «the Little Prince» during the whole-brain fMRI recording. The data is preprocessed by Christophe Pallier with ME-ICA pipeline (Kundu, Inati, Evans, Luh, & Bandettini, 2012).

Measures

Embedding validation. For conceptual-similarity, classic and decorrelated word-embedding models, the pearson and spearman correlation tests will give scores of semantic similarity and association.

Embedding-to-BOLD regression. Regression from each embedding scheme to individual fMRI data will give a correlation of determination (R^2 -value) for each voxel. We will compare the R^2 -value of each embedding model and build a voxel-wise activation profile map as similar in (Jain & Huth, 2018). This would allow us to discover if there is a conceptual-similarity based semantic representation in the previously discovered semantic hub.

Predictions

If the conceptual-similarity space is well built, we expect it to give significantly above null results over similarity benchmarks, and near null results over association benchmarks.

If the dissociation algorithm works as expected, the dissociated association embedding space, would have significantly lower performance on similarity task when compared with conceptual-similarity embedding space, and have comparably similar performance on association task when compared with the undissociated original mixed embedding space.

If transmodal hubs store pure conceptual, transmodal information hierarchically, and other functional neural networks encode other information, then ontologies such as WordNet (Miller, 1995) is analogical to hubs organizational structure. ROIs, which have a preference for conceptual-similarity based embedding models such as WordNet embeddings, should compose an transmodal semantic hub near ATL. Other regions significantly encoded by classic word-embeddings models should have a preference for decorrelated association embedding space.

Analyses

Each built embedding space is tested on semantic similarity and association benchmarks with pearson and spearman correlation. Inter-embedding-space benchmark result comparisons would be tested for significance.

Embeddings with adequate performance in either similarity or association domain would then be used to encode fMRI BOLD signals. The R^2 -values will be tested for significance. We subtract obtained R^2 -values of conceptual-similarity model from association model to build a contrast map with a comparison significance mask. We will run an ANOVA on all voxel R^2 -value between-model differences with subject, embedding type as factors. The voxels with significant main effect of embedding type would draw an additional contour on the contrast map to help determine the localization of a semantic hub graphically.

Interpretation

If the hypothesis is correct and our assumptions on data manipulations are exact, we should see activation preference for conceptual-similarity embeddings in brain regions near ATL,

centered on vIATL. Other significantly correlated voxels found by classic word-embedding models should have a preference for association embeddings. This would further suggest the hierarchical concept organization in the transmodal semantic hub.

Expected contributions

Songsheng Ying. Word-Embedding preparation, embedding space decorrelation, fMRI data analysis and interpretation, master thesis.

Sabine Ploux. Result analysis and linguistic interpretation.

Christophe Pallier. fMRI data acquisition and preprocessing, neuro-linguistic interpretation, fMRI encoding regression scripts.

Laurent Bonnasse-Gahot. Embedding space decorrelation.

Bibliography

- Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Pașca, M., & Soroa, A. (2009). A Study on Similarity and Relatedness Using Distributional and WordNet-based Approaches. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 19–27). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from <http://dl.acm.org/citation.cfm?id=1620754.1620758>
- Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016). Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding. *NeuroImage*, 143, 128–140.
<https://doi.org/10.1016/j.neuroimage.2016.08.068>
- Halawi, G., Dror, G., Gabrilovich, E., & Koren, Y. (n.d.). *Large-Scale Learning of Word Relatedness with Constraints*.
- Hill, F., Reichart, R., & Korhonen, A. (2015). SimLex-999: Evaluating Semantic Models With (Genuine) Similarity Estimation. *Computational Linguistics*, 41(4), 665–695.
https://doi.org/10.1162/COLI_a_00237
- Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458. <https://doi.org/10.1038/nature17637>
- Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, 76(6), 1210–1224.
<https://doi.org/10.1016/j.neuron.2012.10.014>
- Jain, S., & Huth, A. (2018). Incorporating Context into Language Encoding Models for fMRI. *BioRxiv*, 327601. <https://doi.org/10.1101/327601>
- Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage*, 60(3), 1759–1770.
- Lofi, C. (2015). Measuring Semantic Similarity and Relatedness with Distributional and Knowledge-based Approaches. *Information and Media Technologies*, 10(3), 493–501. <https://doi.org/10.11185/imt.10.493>
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*,

38(11), 39–41.

- Peelen, M. V., & Caramazza, A. (2012). Conceptual object representations in human anterior temporal cortex. *Journal of Neuroscience*, 32(45), 15728–15736.
- Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., ... Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(1), 963. <https://doi.org/10.1038/s41467-018-03068-4>
- Ploux, S., & Ji, H. (2003). A Model for Matching Semantic Maps between Languages (French/English, English/French). *Computational Linguistics*, 29(2), 155–178. <https://doi.org/10.1162/089120103322145298>
- Pradet, Q., De Chalendar, G., & Baguenier-Desormeaux, J. (2014). Wonef, an improved, expanded and evaluated automatic french translation of wordnet. In *Proceedings of the Seventh Global Wordnet Conference* (pp. 32–39).
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Rubenstein, H., & Goodenough, J. B. (1965). Contextual Correlates of Synonymy. *Commun. ACM*, 8(10), 627–633. <https://doi.org/10.1145/365628.365657>
- Saedi, C., Branco, A., António Rodrigues, J., & Silva, J. (2018). WordNet Embeddings. In *Proceedings of The Third Workshop on Representation Learning for NLP* (pp. 122–131). Melbourne, Australia: Association for Computational Linguistics. Retrieved from <http://www.aclweb.org/anthology/W18-3016>
- Todorovic, S., & Pallier, C. (2018). *Analyses IRMf lors de l'écoute de texte naturel*.
- Verdier, A., Lakretz, Y., & Pallier, C. (2018). *Encodage d'activité neuronale à partir de réseaux LSTM* (pp. 1–21). Saclay: Neurospin.

Contents

Declaration of Originality	iii
Declaration of Contribution	v
Preregistration	vi
1 Introduction	1
1.1 Semantic Memory, Representation and Processing	1
1.2 Syntagmatic and Paradigmatic Axes in Linguistics	2
1.3 Computational Semantic Representation Modeling	3
1.3.1 Symbolic Relational Semantic Models	3
1.3.2 Statistical Distributed Representations	4
1.4 Semantic Neural Encoding/Decoding Experiments	5
1.4.1 Multi-Network Participation in Semantic Processing	6
1.5 Outline	7
2 Hypotheses	9
2.1 Reconciliation of Multiple Theories Exhibiting the Twofold Character of Natural Languages	9
2.1.1 Semantic <i>Similarity</i> and Semantic Hub	9
2.1.2 Semantic <i>Association</i>	10
2.1.3 Approximative Structure of Twofold Characters in Statistical Distributional Representations	10
2.2 Targeted Semantic Hub Locus	10
3 Methods	13
3.1 fMRI Acquisition and Preprocessing	13
3.2 Semantic Feature Embedding Construction	14
3.2.1 Semantic Similarity Embedding	14
3.2.2 Semantic Association Embedding	15
3.2.3 Embedding Validation	17
3.3 fMRI Voxel-Wise Encoding	17
3.3.1 fMRI Textual Stimuli Preparation	18
3.3.2 Regression Feature Generation	18
3.3.3 Feature Selection for Specific Corpus	19
3.3.4 Ridge Regression with Step-wise Forward Feature Selection, Grid Search and Cross Validation	20
3.4 Analysis	21
3.4.1 Incremental Nested Model Sequence	21
3.4.2 Embedding Contrasts	23
3.4.3 ROI-level Analysis	23
4 Results	25

4.1 Semantic Embeddings	25
4.1.1 Validation on English Data	25
4.1.2 Application on French data	25
4.2 Computational Analysis of Ridge Regression	27
4.2.1 Regressor Generation	27
4.2.2 Choice of α and Effective Feature Dimensionality	29
4.3 Cognitive Analysis of fMRI Encoding	31
4.3.1 Non Semantic-Embedding Models	31
4.3.2 <i>Similarity</i> Nested Model	34
4.3.3 <i>Association</i> Nested Model	36
4.3.4 <i>Similarity/Association</i> Contrast	37
5 Discussion	41
5.1 Back to Hypothesis	41
5.2 Precise and Informative Semantic Feature Design	41
5.3 Limits of fMRI	43
5.4 Statistics	44
5.5 Cognitive Accounts on Coherence between Embeddings, Semantic Principles and Semantic Hub	44
6 Conclusion	47
A Supplementary Methods	49
A.1 fMRI Stimuli Preparation	49
A.1.1 Natural Story Stimuli	49
A.1.2 Behavior Control	49
A.2 fMRI Acquisition	50
A.3 Regression Parameters	51
A.4 Supplementary Analysis	51
A.4.1 Non-nested Model Comparison	51
A.4.2 Comprehensive ROI List	52
B Supplementary Results	53
B.1 Semantic Embeddings	53
B.1.1 Visualization of Semantic Spaces with TensorFlow Projector . .	53
B.1.2 Semantic Ranking Task Results	53
B.1.3 Vocabulary Coverage by POS	53
B.1.4 Corpus-Targeted Semantic Feature Selection	53
B.2 Non-nested Model Comparison	53
B.3 Regression	59
B.3.1 More on α and Dimension Selection	59
B.3.2 Impact of CWRATE	60
B.4 Embedding Model Brain Maps	60
B.4.1 SIM and ASN F-test	60
B.4.2 SIG Nested Model and Contrast	60
B.5 On Behavioral Control	66
Bibliography	69

List of Figures

1.1	Impact of Context Window Size on Syntagmatic and Paradigmatic Information Extraction	5
3.1	Explained Variance Ratio of WordNet Embedding Principle Components	15
3.2	Smoothed Differentiated EVR of WordNet Embedding PCs	16
3.3	SPM Hemodynamic Function	18
4.1	EVR of 4 Semantic Spaces, English	26
4.2	EVR of 4 Semantic Spaces, French	27
4.3	French SIM Regressor Variances	30
4.4	Subject Best Hyper-parameter Configuration Voxel-Count Heat-map .	30
4.5	Encoding with BASE Features, Group	31
4.6	Histogram of r^2 with BASE Features	33
4.7	Encoding with SIM Features, Group	34
4.8	Histogram of r^2 with SIM/ASN Features	35
4.9	Encoding with ASN Features, Group	36
4.10	SIM-ASN Contrast, Group	37
4.11	SIM-ASN ROI Contrast, Group	38
B.1	French SIM Space Visualization	55
B.2	French ASN Space Visualization	55
B.3	English SIM Space Visualization	55
B.4	SIM Predicted ASN Column Model	58
B.5	ASN Predicted SIM Column Model	58
B.6	Typical-Voxels' Response to α and Dimension	59
B.7	Session Best Hyper-parameter Configuration Voxel-Count Heat-map .	60
B.8	Best Hyper-parameter Configuration Voxel-Count Histograms	61
B.9	Encoding with SIG Features, Group	63
B.10	SIG-ASN Contrast, Group	64
B.11	SIG ASN ROI Contrast, Group	66

List of Tables

1.1	Involvement of Cerebral Areas in Semantic Tasks	1
1.2	Example of Syntagmatic and Paradigmatic Axes	2
1.3	Example of Syntagmatic and Paradigmatic Mixture in Statistical Semantic Models	4
4.1	English Semantic Space Semantic Ranking Task Results	26
4.2	French Semantic Space Semantic Ranking Task Results	28
4.3	The Little Prince Vocabulary Coverage in Semantic Spaces	29
4.4	RMS/CWRATE/SIM/SIG/ASN Best Modeled Voxel Clusters	32
4.5	CWRATE Voxel Improvement Clusters	33
4.6	SIM Voxel Improvement Clusters	35
4.7	SIM-ASN Voxel Contrast, SIM, Group	38
4.8	SIM-ASN Voxel Contrast, ASN, Group	39
A.1	French <i>The little prince : a French/English bilingual book</i> Chapter Division	49
B.1	Exemplar Neighborhoods in French Semantic Embeddings	54
B.2	English WordNetEmbedding Iterations	56
B.3	The Little Prince Vocabulary Coverage by POS	57
B.4	SIM Predicted ASN Column Model Performances	58
B.5	ASN Predicted SIM Column Model Performances	58
B.6	ASN Voxel Improvement Clusters	61
B.7	CWRATE/SIM/SIG/ASN F-test Significant Voxels	62
B.8	SIG Voxel Improvement Clusters	63
B.9	SIG-ASN Voxel Contrast, SIG, Group	65
B.10	SIG-ASN Voxel Contrast, ASN, Group	66
B.11	Model Performance and Behavioral Question Score Correlation	67

List of Abbreviations

BA	Broadmann Area
BOLD	Blood-Oxygen-Level Dependent
CSC	Controlled Semantic Cognition
CV	Cross Validation
EEG	electroencephalography
ERP	Event Related Potential
fMRI	functional Magnetic Resonance Imaging
GLM	General Linear Model
MEG	magnetoencephalography
PC	Principle Component
PET	Positron Emission Tomography
POS	Part-Of-Speech
ROI	Region of Interest
SD	Semantic Dementia
SDR	Statistical Distributed Representation Models

Embedding Spaces

SIM	SIM ilarity
SIG	SIM ilarity projected on DepGlove
ASN	AS sociatio N
MIX	MIX ed

List of Cortical Region Name Abbreviations

*L	* Lobe
*G	* Gyrus
*S	* Sulcus
S*	Superior *
M*	Middle *
I*	Inferior *
IFG	Inferior Frontal Gyrus
IFGoper	... pars opercularis
IFGorb	... pars orbitalis
IFGtri	... pars triangularis
IPG	Inferior Parietal Lobe
SPG	Superior Parietal Gyrus
AG	Angular Gyrus
PFC	Prefrontal Cortex
SMG	Supramarginal Gyrus
a*	anterior *
m*	middle *
p*	posterior *
*TL	Temporal Lobe
... aTL	anterior TL
*TG	Temporal Gyrus
STG	Superior TG
MTG	Middle TG
*TP	Temporal Pole
TPJ	Temporoparietal Junction

Chapter 1

Introduction

1.1 Semantic Memory, Representation and Processing

Semantics as in linguistic context, is the connection between language forms such as orthography, syntax, and meanings including lexical and phrasal ones. The brain processes semantics in aide of semantic memory (Tulving, 1972), of which the loci in human cognition system are still actively being debated.

Semantic memory is unlike episodic memory, which is individual-specific and modality dependent. As semantic memory is associated with general world knowledge (McRae & Jones, 2013), semantic memory tends to be shared across individuals within a common cultural background. Such invariance provides a window for semantic memory loci localization in the human brain.

In Tulving's view, semantic memory is conceptually dissociated, but not necessarily functionally or structurally separated from procedural and episodic memory. However, studies (Vargha-Khadem et al., 1997) suggest the structural and functional dissociation of episodic memory and semantic memory, and that parahippocamal cortices, as a classical theoretical locus of episodic memory, are not crucial for the normal functioning of semantic memory.

While some theorist argue for the temporal localization of semantic memory (Martin & Chao, 2001; Saumier & Chertkow, 2002), recent studies found neural correlates of semantic knowledge distributed in multiple lobes (refer to Table 1.1 for a resume), supporting thus the distributed hypothesis that semantic knowledge might be encoded in multiple brain areas. The cortices involved in semantic processing other than the temporal cortex, are strongly associated to perceptual, sensorial

Reference	Frontal Lobe	Temporal Lobe	Parietal Lobe	Occipital Lobe	Limbic Lobe
Tsukiura, Mochizuki-Kawai, and Fujii, 2006 [TODO More weekend]	IFG	CA, STG	AG	GF	PCG

TABLE 1.1: Involvement of Cerebral Areas in Semantic Tasks

syntagmatic						
paradigmatic	The	ridiculous girl	fell	into	the	pond.
	silly	person	jumped			river.
	foolish	woman	tripped			lake.
	funny	lady	plunged			sea.
	crazy	princess	walked			ocean.
	klutzy	child	ran			pool.

TABLE 1.2: An example of syntagmatic and paradigmatic axes. Gray-colored texts are in *absentia*, black-colored texts are in *presentia*. *Syntagm* combines word sequence into a meaningful sentence, while *paradigm* provides feasible substitutions of currently-present words.

and/or affectual functions, suggesting that semantic memory depends on episodic events (Moseley & Pulvermüller, 2014).

A theoretical reconciliation between these two schools is the specialization of cortical areas implicated in semantic tasks: the abstract, amodal semantic memory being grounded by concrete modal episodic memories (Pecher & Zwaan, 2005).

The classical semantic memory locus could serve as the convergence zone for binding information from modality-specific cortices (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Simmons & Barsalou, 2003). Anatomical evidences, that the possible loci sit at convergences of multiple perceptual processing streams (Binder & Desai, 2011), support this theory. Patterson, Nestor, and Rogers (2007) further exploited Damasio et al.'s arguments by proposing the *Hub-and-Spoke* model, where a semantic hub, not only acts as a pointer and a information-binder, but also constructs, refines semantic concepts and builds cross-modal similarity structures using episodic event. In parallel, Paivio (2008) argued for a dual-coding system to address the problem of representing abstract concepts which do not necessarily have perceptual input: in addition to accumulated perceptual information for concrete concepts, a more meta-semantic department keeps record for all concepts.

In this project, we restrain the discussion to lexicosemantic system, particularly we will focus on the representation and processing of word meanings in an ecological auditory experiment.

1.2 Syntagmatic and Paradigmatic Axes in Linguistics

Jakobson and Halle (1963) and De Saussure (1969) propose that all linguistic units are arranged in two modes which are *combination* and *selection*, or *syntagm* and *paradigm*. *Combination* is in *presentia* as the linguistic unit (in the current context, a word / lemma / lexicon unit) is contextualized by other elements presented in a linguistic sequence. *Selection* is in *absentia* as it is linked to other alternative substitutions which are absent from the current context. Table 1.2 gives an example of the organization in two axes.

Jakobson and Halle further illustrated the twofold character of language via selection-deficient and contexture-deficient aphasics using data from Crutch and Warrington (2004), Goldstein (1948, 1971), Head (1920), Hughlings Jackson (1879), Luria (1976),

Warrington and McCarthy (1983), bridging formal linguistic works with psycholinguistic studies.

Similarity disorder (selection-deficiency) patients are able to complete scraps of words or sentence, but are unable to uncontextualize themselves, thus unable to utter "it rains" unless it rains actually. The retrieval of the most precise lexicon is blocked, and the internal relation between concepts are dissolved. For those patients, an isolated word means nothing, occurrences of one same word in different contexts are homonyms, and the production of word tends to be bound by other associative words (for example, *knife* are referred to as *pencil-sharpener*, *bread-knife* and *knife-and-fork*) or metonymies (*fork* for *knife*, *eat* for *toaster*), or replaced by the most general terms such as *chose* and *machin* in French. The utterances are highly dominated by spatial, temporal and usage proximities, which are not necessarily parallel to similarity. They also lose the ability to switch register and stay in their idiolect reality. As remarked by Jakobson and Halle, for such an aphasic whose substitutional capacity has been disabled and contextual capacity intact, the emissive and receptive linguistic competence relies solely on contiguity.

Contiguity disorder (contexture-deficiency) patients, on the other hand, are impaired to propositionize, inflect and desolve compound words such as *thanksgiving* into *thanks* and *giving*. They produce agrammatical sentences as a chaotic word heap. The approximative identifications of a presented concept are quasimetaphoric (such as *spyglass* are produced for *microscope*, *fire* for *gaslight*), without any deliberate transfer of meaning as it is in the case of poetry and rhetorics.

1.3 Computational Semantic Representation Modeling

Natural language processing and understanding in general artificial intelligence has partially branched away from cognitive computational linguistic works. While language representation models like BERT (J. Devlin, Chang, Lee, & Toutanova, 2018) are fine-tuned to natural language processing benchmark tasks, they do not necessarily approach human language processing. We restrain computational semantic modeling to the models attempting to replicate of human linguistic dynamics.

Semantic representation models digitalize the natural language word meanings into numeric representations that can be understood and processed by neural networks and computer systems. There are two schools of modeling: symbolic and distributional.

1.3.1 Symbolic Relational Semantic Models

Classical semantic models assumes the meanings can be considered as an indexable binary feature array (Smith, Rips, & Shoben, 1974) or interconnected nodes in a large semantic graph-like ontological network (Collins & Quillian, 1969). In such structures, the binary features and nodes in the ontologies each represents a semantic entity (*symbol*), to which we associate properties or values. Depending on the implementation, such symbolic structures are usually abstracted or independent from episodic, perceptual experiences. They are able to account for abstract taxonomical conceptual comparisons. Therefore, they model mainly paradigmatic similarities.

Modern implementation of such models still rely largely on human manual coding. WordNet-alike knowledge bases (Miller, 1995, 1998; Pradet, De Chalendar, &

Target word:		teacher	
Neighbour	Cosine Distance ¹	Nature of Neighbour	Semantic Relation
classroom	0.537	associate	locative
teaching	0.497	associate	action
school	0.484	associate	locative
preschool	0.453	associate	locative
student	0.421	associate	object/agent
grade	0.418	associate	
college	0.403	associate	locative
instructor	0.401	synonym	

* A cosine distance near 0 indicates a greater similarity.

TABLE 1.3: *Teacher* are judged to be close to *classroom*, *teaching*, *student*.... While they are frequent collocations, they are nevertheless not synonyms.

Baguenier-Desormeaux, 2014; Sagot & Fier, 2008) are examples of symbolic semantic networks which encodes inter-word semantic relations. In this class of models, lexicon units are regrouped into *synsets*, forming synonymy sets, each representing one different meaning of the unit. Synsets are interconnected with relations such as *antonymy*, *hyponymy*, *hypernymy*, *meronymy*, *toponymy*...

1.3.2 Statistical Distributed Representations

Harris (1954)'s distributional hypothesis argues that "linguistics items with similar distributions have similar meanings." Most statistical models based on this theoretical foundation could be classified into latent semantic inference models (United States Patent, 1989; Pennington, Socher, & Manning, 2014) and hyperspace analogue to language models (Burgess & Lund, 1995; Levy & Goldberg, 2014; Mikolov, Chen, Corrado, & Dean, 2013). As they make heavy use of contextual information, the syntagmatic information are also present those classes of models.

Such representation models use high dimensional vectors to encode semantic entities. The 2D matrix representation of the model, where the rows are entries of the lexicon, columns being the vector dimensions, are referred to as *semantic embeddings* or *semantic spaces*. Similarity measures are derived on vector distance metrics including cosine distance, gaussian distance and Minkowski distances. Models such as Mikolov et al. (2013), Pennington et al. (2014) successfully capture semantic information from textual statistics, achieving adequate performance on similarity benchmarks.

On the linguistic nature of statistical distributed representation (SDR) models, we observed a mixture of syntagmatic and paradigmatic information in statistical embeddings. To give an example, in an openly available GloVe (Pennington et al., 2014) implementation¹, the closest neighbors of the target word *teacher* (a noun) are composed of synonyms (*instructor*, *tutee*) and associates (*classroom*, *teaching*, *school*, *student*, *aunt*...) (Table 1.3). While the list of synonyms proposed by WordNet is *instructor*, *teaching fellow*, *docent*, *coach*, which is purely paradigmatic.

¹Wikipedia 2014 + Gigaword 5 with 6B tokens, 400k uncased vocabulary and 300 dimensions. <https://nlp.stanford.edu/projects/glove/>.

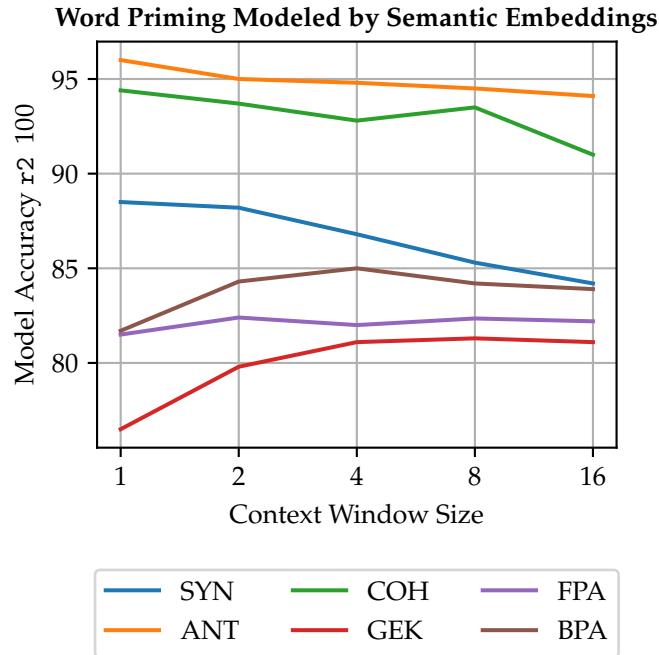


FIGURE 1.1: Lapesa, Evert, and Schulte im Walde (2014) tested 6 priming scheme datasets: paradigmatic datasets include synonyms SYN, antonyms ANT and cohyponyms COH, syntagmatic datasets include forward phrasal associates FPA, backward phrasal associates BPA and generalized event knowledge GEK. For each of the 6 datasets, they trained a separate GLM using differently configured semantic embeddings to predict word priming delays. Increasing context window size when training the embedding improves syntagmatic model accuracy, while penalizes paradigmatic predictions. Note that paradigmatic results are consistently better than syntagmatic ones.

Lapesa, Evert, and Schulte im Walde (2014) tested combinations of different hyperparameters of co-occurrence-based statistical representation building algorithm. They used behavioral priming data of syntagmatic and paradigmatic word-pairs to contrast parameters' influence on two axes' performance. Figure 1.1 is reproduced using the reported data from the work, confirming more systematically the two-fold mixture in SDR embeddings.

1.4 Semantic Neural Encoding/Decoding Experiments

The cognitive account of semantic processing includes the identification of the specific functions of various cortical areas during the semantic process. Historically, neuroscientists had to rely on semantic deficits and lesion studies. With cognitive modeling development, neuroimaging techniques enabled the examination of various model proposals without having to open the skull.

Marr (1982) proposed the three levels of modeling in: *computation*, *algorithm* and *implementation*. On operating on the computational level, cognitive modelers either try to replicate the temporal-spatial dynamics of cerebral activities, thus *encodes* neural signals, and/or use measured signals to recover external stimuli, thus *decodes* brain activity.

Given the temporal and spatial resolution constraints of neuroimaging techniques among electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), fMRI and PET gained enormous popularity in the semantic localization studies while MEG for spatial-precise temporal dynamic studies (Mollo, Cornelissen, Millman, Ellis, & Jeffries, 2017). fMRI gives a much better spatial resolution than EEG (up to millimeters), but it is generally poor in temporal resolution. Classic semantic stimuli units are presented at sub-second level (hundreds of milliseconds) but the usual imaging frequency is at second level, and the measured blood-oxygen-level dependent (BOLD) signal approximates a temporal convolution of real neuron activation and hemodynamic function over 4 to 7 seconds.

Encoding and decoding experiments heavily rely on computational semantic representation models, especially distributed representations. Mitchell et al. (2008) was the first work to use contextual co-occurrence vectors (one variate of SDR) to encode semantic processing activities for concrete nouns. The predictive model showed significant generalization power, indicating a strong association between semantic embeddings and the brain activity, and the feasibility of encoding fMRI recorded semantic brain activity.

1.4.1 Multi-Network Participation in Semantic Processing

A large amount of literatures with encoding and decoding experiments report results in favor of distributed semantic processing in the human brain.

Mitchell et al. (2008)'s concrete noun encoding experiments reported the most accurately predicted voxels to be located in (pre)frontal, parieto-temporal regions. Pereira et al. (2018) used fMRI signals to decode lexical and phrasal semantic stimuli presented in isolation. Among the 5000 most informative voxels found for each subject, functional networks in human brain other than the linguistic one take also a significant portion consistently across subjects.

In addition to experiments conducted with isolated semantic stimuli, works such as Todorovic and Pallier (2018), Verdier, Lakretz, and Pallier (2018) [TODO, add more refs] experimented with ecological stimuli. They found consistent encoding / decoding performance with non-classically-linguistic cortical area voxels.

Multiple theories exist to account for the contribution of other cortical areas in semantic processing.

Evidences for Feature-Based Distribution

Feature-based distributional models argue that various cortices are recruited to encode modality-specific information (Chao, Haxby, & Martin, 1999; Goldberg, Perfetti, & Schneider, 2006; Hauk, Johnsrude, & Pulvermüller, 2004). The parallel activations of these areas participate in the completion of semantic retrieval and representation (Patterson et al., 2007). Experiments (Borghesani et al., 2016) successfully decode perceptual information (size, [TODO] ...) associated with isolated words using only textual stimuli. Huth, De Heer, Griffiths, Theunissen, and Gallant (2016) proposed a semantic cortical mapping organized by PCA-generated semantic axes including perceptual properties (e.g. visual, tactical, emotional, locational) along with semantic domains(e.g. tools, animals, living animates). Rowtula, Oota, Gupta,

and Surampudi (2018) mixed textual semantic embeddings with image-based (visual) semantic embeddings to encode Pereira et al. (2018)'s data. The multimodal model, compared to purely textual ones, gave much better results in fMRI encoding.

Evidences for Semantic-Domain-Based Distribution

Domain-specific distributional models (Damasio et al., 1996; Damasio et al., 2004; Mahon & Caramazza, 2011) argue for a cortical map in function of semantic categories (*domains*, such as living animate, vegetables, tools). The argument is mainly supported by category-specific pathology observations: cortical connectivity are locally tuned for different semantic topics' operational processing.

Huth, Nishimoto, Vu, and Gallant (2012) used WordNet-based noun and verb hierarchical structures to correlate neural responses in different cortical area with word categories (*semantic domain*, e.g. *athlete*, *communicate*). Analysis showed that voxels in [TODO replace labels by understandable names: IFSFP, FO, AC, FFA, PPA, COS, ITS, V1-V7, OFA, LO, IPS, RSC, MPC ...] responded to different domain-specific words. The semantic axes of Huth et al. (2016) mentioned in Section 1.4.1 include also domain-specific ones.

Importantly, semantic features and domains are not necessarily two dissociated principles of semantic organization. For example, domain traits can also imply *feature* information (domestic animals imply the size of the concept in question shall normally not surpass that of an adult human.) [TODO: better example]

Evidences for Semantic Control Networks

Controlled Semantic Cognition (Ralph, Jefferies, Patterson, & Rogers, 2017) system argues for an operational rather than representational account. CSC argues for a semantic hub(Patterson et al., 2007), and it considers the neural correlates in non-hub areas as the interaction with semantic representation system and the computation of semantic entities and non-linguistic decision making (Fuster, 2004). Semantic computation, such as combination and selection, is modulated by linguistic and task contexts. Since language is also a social tool, goal, action and decision making is also implicated in semantic processing. These functions recruit the cortices revealed by semantic encoding/decoding experiments.

1.5 Outline

This master's project attempts to examine different cortical regions' participation in the two semantic processing axes, which we name as *similarity* and *association*. We attribute each axis, or organisational principle, with a fine-tuned semantic representation model, configured with our assumptions and hypothetical properties on these axes. Through an fMRI encoding experiment using lexical semantic models in an ecological setting, we infer each region's functional properties based on the locally preferred organisational principle.

In chapter 2, assumptions on *similarity* and *association* axes and semantic distributional representation structures are presented and discussed along with the semantic memory localization hypotheses. In chapter 3, we will present our methods on building semantic representations, preliminary assumption validations, fMRI encoding settings and result interpretation. Then the obtained results are presented in

chapter 4. We will present further ad-hoc analyses and possible implications of our results in chapter 5.

Chapter 2

Hypotheses

We relate De Saussure and Jakobson and Halle's twofold structuralism to neuropsycholinguistic theories on semantic processing. The paradigmatic axis is associated with semantic *similarity* relations, and syntagmatic axis with semantic control or episodic-event based semantic information comparisons, thus *association* relations. Two parallel (not necessarily separate) systems handle respectively metaphorical and associational retrieval and access of words in linguistic tasks.

2.1 Reconciliation of Multiple Theories Exhibiting the Twofold Character of Natural Languages

In continuity with Jakobson and Halle (1963)'s syntagmatic and paradigmatic axes and multiple theories on the role of cortices which are not traditionally defined as language areas, we focus on dissecting a central semantic locus, which acts as the binder, gateway or hub in different theories' nomination, apart from other peripheral semantic components.

2.1.1 Semantic Similarity and Semantic Hub

Semantic hub is the locus/loci where the ontological semantic information of all words can be accessed. Such a particular ontology encodes the human understanding of concepts free of the dominant influence of modality-specific semantics. The plausibility of such a hub is motivated by contexture-deficients' quasimetaphoric wordings and selection-deficients' uncontextualizability (Section 1.2). Patterson et al. (2007) also summarized *semantic dementia* (SD) symptoms including concept retrieval and categorization difficulties.

Evidences from SD studies indicate three principle factors in semantic hub organization: familiarity, typicality and specificity. Familiarity is constructed with episodic events (thus out of scope of this project). Typicality can be encoded in the semantic ontology since untypical concepts usually require more information (e.g. *whale* is conceptually very similar to other marine fishes, thus it needs to be marked as an exception in the semantic system since it is a mammal). Specificity can be modeled by a hierarchical structure, where the access to a word is an iterated traverse in a semantic tree. Hierarchical structures also allows to cope with a large lexicon inventory. Such a hypothetical construction can reproduce the paradigmatic axis, since semantic hub implementations should be able to group and cluster animal names such as cat, tiger, dinosaur together locally, location names such as museum, zoo, schools and color names such as magenta, grey, yellow.

The internal organization of word-meanings in the semantic hub are henceforth named *similarity*. It has a more global view towards all word meanings, whereas paradigmatic relations are a local similarity manifestation.

For computational implementation of *similarity*, we assume the amodality and potential hierarchy of semantic *similarity* is well conserved in ontological semantic networks introduced in Section 1.3.1. WordNet-like networks hand-code its semantic units, introducing thus a familiarity bias. Furthermore, they explicit various semantic relationships: hypernymy and hyponymy are considered as the backbone structure of similarity hierarchy, synonymy (formed by *synsets*) pushes similar words closer...

2.1.2 Semantic Association

We consider *association* as complementary to *similarity* in semantic memory. It includes thus syntagmatic relations, modality-specific proximities and episodic associations. Extended syntagmatic relations include collocations (*pencil* and *write*), meronymy/holonymy (*ceil* and *house*), entailment/causality (*sunset* and *milkway*). Modality-specific proximities include spatial proximities (*bridge* and *river*), visual similarity (Paris metro logo and McDonald's), rhymes (*rhyme* and *lime*)...

Domain-specific information can also be accounted such as *tools* recruit sensory-motor functions, *music instruments* require auditory functions, and *human faces* calls for affectual and social departments.

Association is vast. In this project we model the *association* axis by exploiting all available information from statistical distributed representations (SDR), in which syntagmatic relations, common episodic associations (manifested in the SDR source corpus) are present.

2.1.3 Approximative Structure of Twofold Characters in Statistical Distributional Representations

As SDR is a mixture of *similarity* and *association* information, we approximate this mixture by a linear additive structure of two components. Despite its simplicity, linear structuralism is often adopted in semantic modeling (e.g. Continuous-Bag-of-Words which is proposed in Mikolov et al. (2013)), and it achieves adequate performance.

With this approximation, we can obtain a semantic *association* representation space by subtracting *similarity* representations from a mixed representation space. Once with the *similarity* component removed, the embedding should rank associates above the residual synonymies.

2.2 Targeted Semantic Hub Locus

With the key question on finding the semantic information encoded by various cortical departments engaged in semantic processing, we propose to test a possible theoretical construction based on the dissociation between *similarity* and *association* axes.

As classical view holds that the temporal cortex hosts semantic memory, more precise semantic hub loci are proposed by different theories. Binder et al. (2011), Patterson et al. (2007), Price, Devlin, Moore, Morton, and Laird (2005) argue for an bilateral

aTL loci, whereas Ralph et al. (2017) refined the search to ventrolateral aTL. As only at a large-scale is the spatial placement of anatomical convergence zone predictable across individuals (Damasio et al., 2004), we ground our hypothetical precision to aTL.

If the so argued semantic hub exists and its loci are stable across individuals, theoretical reconstruction of cerebral activity based on semantic *similarity* embeddings should better model semantic hub region activities namely aTL, whereas *association* embedding reconstructions in other language areas.

Chapter 3

Methods

To address of problem of localizing cortical areas processing semantic *similarity* and / or *association* , we use voxel-wise fMRI encoding to estimate local superiority of semantic representations. The human subjects' brain activity are recorded when they attentively listen to naturally spoken narrative stories from *Le Petit Prince*. We construct features (regressors) using non-semantic signals (including acoustic energy, word presence and content word presence) and semantic representation models tuned for semantic *similarity* and *association* .

3.1 fMRI Acquisition and Preprocessing

The fMRI experiment is designed and carried out by Todorovic and Pallier (2018). 20 French native speakers (11 females, average age of 24.5 years-old, range 18-39 years-old, right handed according Edinburgh's inventory (Oldfield, 1971) (adapted for French, averaged score 0.903, range 0.375-1), without antecedent neurological or psychiatric disorders) were recruited from Neurospin's volunteer inventory. The participants listened to the French audio book *The little prince : a French/English bilingual book* (Antoine de Saint-Exupéry & Wilkinson, 2011) during 9 runs. They were tested with comprehension multi-choice questions at the end of each block. During the listening period, a Siemens scanner scanned the whole brain at 3 Tesla with multi-echo EPI sequence at 2 second-per-image rate. Each scanning session (run) lasted at maximum 90 minutes. Each subject passed all the recording during the same day.

The multi-echo procedure has a higher signal-to-noise ratio over mono-echo. Therefore, activities in traditionally unaccessible cortical areas such as the ventral temporal cortex. The voxel size, as a compromise, is fixed at a larger volume than classic modern fMRI recordings of $3.159 \times 3.159 \times 3.159 mm^3$. The acquired MRI (anatomical and functional) data are then preprocessed with ME-ICA pipeline¹ (Kundu, Inati, Evans, Luh, & Bandettini, 2012) to transform spatial normalization to MNI template and extract whole-brain BOLD signals.

Please refer to appendix Section A.1 for comprehension question designs and fMRI stimuli preprocessing, Section A.2 for more detailed participant recruitment, fMRI procedure presentations, and to the original report (in French) (Todorovic & Pallier, 2018) for original materials used in the experiment.

¹Library available at <https://github.com/ME-ICA/me-ica>, commit 6ae63c7.

3.2 Semantic Feature Embedding Construction

To build regressors for voxel models of different semantic processing axes, we first obtain corresponding semantic embedding of the semantic principles in question. We then validate the embedding performances in *similarity* and *association* semantic proximity ranking evaluations before being used to build fMRI regressors.

Since we do not have widely-used French evaluation task benchmarks in disposition, our implementation of semantic embedding algorithms are tuned and validated firstly on English data, then transferred on French data.

3.2.1 Semantic Similarity Embedding

To build semantic similarity representation, English WordNet² (Miller, 1995, 1998), French WOLF (Sagot & Fier, 2008) serve as our data source. For semantic entities encoded in an ontological graph with internal semantic relations, Saedi, Branco, António Rodrigues, and Silva (2018) proposed an evaluation of semantic affinity by counting all the paths connecting two nodes representing entities. The paths are indexed by semantic relationships, and are weighted by their length: shorter is the path, stronger is the semantic affinity. Equation 3.1 illustrates the exact numerical calculation by taking M as the weighted adjacency matrix representation of the initial graph.

$$M_G^{(n)} = I + \alpha M + \alpha^2 M^2 + \cdots + \alpha^n M^n M_G = \sum_{e=0}^{\infty} (\alpha M)^e = (I - \alpha M)^{-1} \quad (3.1)$$

After computing the graph distance of between all word-pairs, a normalized Positive Point-wise Mutual Information transformation is applied to the matrix to reduce noises introduced by unbalanced word occurrence frequency. Finally a PCA is performed to reduce the dimensionality of the large matrix.

We replicated Saedi et al.'s experiment using the reported optimal parameters³. Our tests differed from the original work's configuration on semantic relation selection, vocabulary size and dimensionality choice of the resulting embedding space.

As mentioned in the Section 2.1.1, typical similarity-related relationships include synonymy, hypernymy, hyponymy. In WordNet implementation, we extend these relationships to include also the relation where an adjective is a participle of verbs (e.g. *exhausting* is a participle of *exhaust*), is similar to another adjective (*exhausting* and *tiring* are similar to *effortful*) and where an adverb is derived from an adjective (*essentially*, *basically* and *fundamentally* are derived from *essential*). We test different set of combinations of relations to further confirm our choice for *similarity* and *association* classification (see Section B.1.2).

As large matrix calculation is very memory-consumptive, given the available memory on the laboratory server⁴, we selected the 15 000 most frequent words in WordNet and 20 000 in WOLF for fast in-memory computing, and then took 60 000 in WordNet and all words in WOLF (56665) for optimal embedding performances.

²Version 3.1, available at <https://wordnet.princeton.edu/download/current-version>

³including graph random-walk decay factor, semantic relation weight attribution

⁴The computer is equipped with an quad-core Intel Xeon processors @ 3.70 GHz, 32 GB RAM, running Ubuntu 18.10, Python 3.6.7 Anaconda.

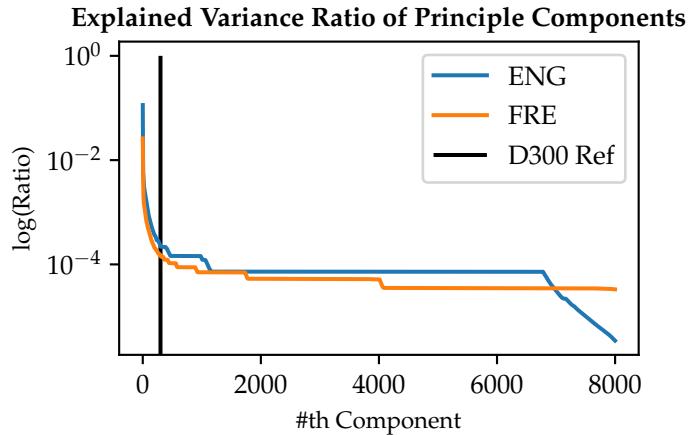


FIGURE 3.1: The log of EVR of each PC in WordNet embedding space (ENG) and in WOLF embedding space (FRE). The PCs are ordered by their corresponding eigenvalues. The black vertical line is placed at the classic choice dimensionality of 300 as a reference.

A balance between model complexity and precision is required. To determine the number of principle components (PC) of the graph encoding word-wise proximity, we initially keep first 8 000 PCs as potential candidate (see Figure 3.1). Thresholding directly the singular values or the explained variance ratios (EVR)⁵ resulted to either too few or too much retained PCs. We use the discrete derivative of the EVR to obtain second-degree information. Since the derivative is extremely noisy but a general decreasing trend could be seen, we apply a Savitzky-Golay filter⁶ of window size 100 and of first order on the discrete derivative of EVR. The cutting position is visually selected around the first local minima with a following sufficiently wide valley of the smoothed signal (Figure 3.2). The derivative as the cutting basis is more sensible to information conservation in embedding spaces.

3.2.2 Semantic Association Embedding

$$M = S.P + A \quad (3.2)$$

Using the linear approximation of *similarity* and *association* information mixture in classical SDRs (refer to Section 2.1.3), we extract *association* representations from classic statistical embeddings with equation 3.2, where M is the mixed semantic representation space, S being the *similarity* based space, P a learned projection matrix projecting the similarity space onto the mixed space, and the residual A being the *association* space. The embedding spaces of interest are S and A , henceforth respectively noted as **SIM**(short for **similarity**) and **ASN** (**association**). The two auxiliary spaces are P and M , noted as **SIG** (**similarity** projected on (Dep)Glove) and **MIX** (**mixed**). The projection weight P is learned via a general linear model (GLM)⁷, with the computational objective to minimize the L-2 norm of A .

⁵EVR is the eigenvalue of the PC divided by the sum of all eigenvalues

⁶Low-degree polynomials are fitted on a subset of data points with least-square error minimization, so that data points can be smoothed without distorting the global trends

⁷https://scikit-learn.org/stable/modules/linear_model.html

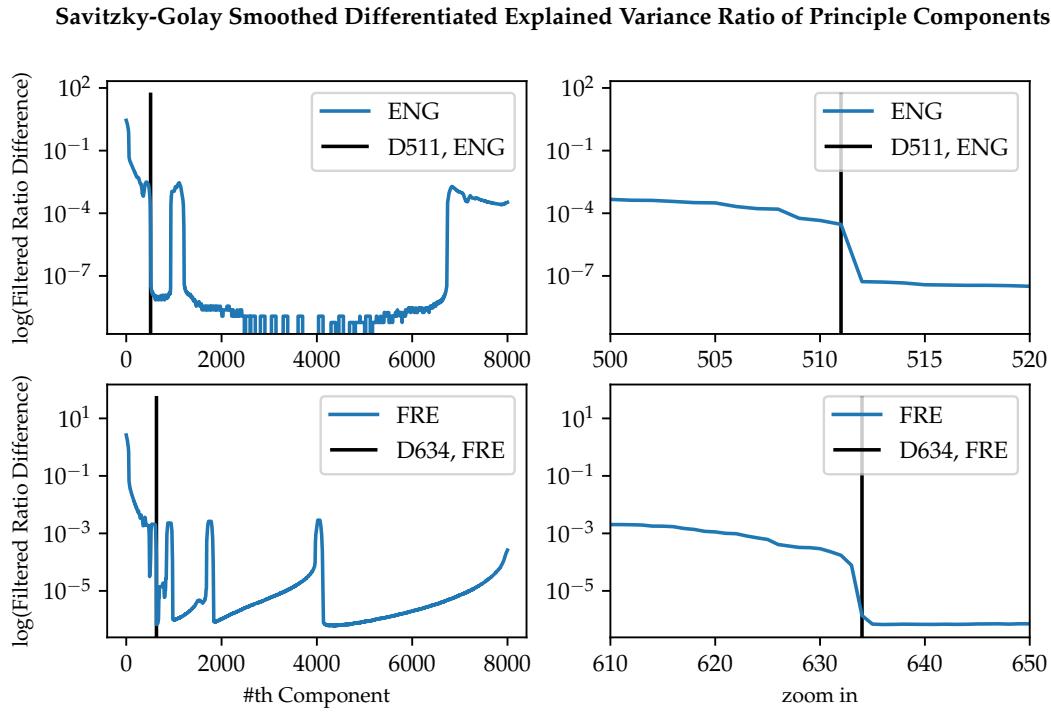


FIGURE 3.2: Left panels: Savitzky-Golay filter smoothed the discrete derivative of the PC EVR signal. D511 in ENG and D634 in FRE are visually selected on the left border of a sufficiently wide signal valley.

For MIX spaces, for English we use GloVe (Pennington et al., 2014) trained on Common Crawl with 840B tokens, 2.2M cased vocabulary and 300-dimension vectors⁸. For French we use DepGloVe⁹ (de la Clergerie, n.d.).

English GloVe embeddings provide word-level vectorial representations. However, since French verbs and adjectives have various inflectional forms, this inflection can thin out captured semantic information if each specific form does not have sufficient frequency in a given corpus. DepGloVe aggregates semantic information by lemmatizing the tokens and attribute them with a part-of-speech (POS) tag. The main POS tags include nc (common nouns), np (proper nouns), v (verb), adj (adjectives), adv (adverbs) along with auxiliary tags.

To formulate the GLM dataset, modifications on heterogeneous embedding data are made to align the embedding spaces. Each row of the an embedding space matrix represents a lexicon unit. Since different semantic spaces have different lexicon settings, only the intersection of two vocabularies of one same language is kept in later stages. The lexicon alignment between the English embeddings is based on orthography, which is computed by string comparison. For French data, multiple text sources are converted to the same format: lemma tagged with WOLF POS tags (nouns, verbs, adjectives and adverbs). We hand coded rules to tidy up WOLF vocabulary and transformed DepGloVe's complex POS tagging entries into WOLF's relatively simple set. Our textual data including validation task benchmarks and fMRI stimuli are also transformed to align with this strategy. Finally we manually check the vocabulary coverage against validation dataset and fMRI stimuli words, we also performed

⁸Pretrained data available at <https://nlp.stanford.edu/projects/glove/>

⁹Pretrained data available at <http://alpage.inria.fr/dep glove/process.pl>

manual correction in the newly aligned space to purge algorithm erroneous results.

3.2.3 Embedding Validation

Since many assumptions and approximations are made on the structure and content of semantic representation spaces, the interpretation of further results depends on the validity of the embedding construction.

The semantic ranking task depends on databases of word-pairs, each attributed with a score (usually annotated by human) which measures the proximity in terms of the semantic property defined by the task. Each semantic embedding is provided with a proximity metric, which could be derived from graph distances or vectorial distances. The score of the task is computed by calculating Pearson's and Spearman's correlation coefficient r between the embedding based word-pair proximity and the baseline.

Conformably with other works on semantic model evaluation methods [TODO Refs, other works using the same benchmarks] (Saedi et al., 2018), we use benchmark data provided by Rubenstein and Goodenough (1965) (**RG1965**), Agirre et al. (2009) (**WS353-Similarity**) and Hill, Reichart, and Korhonen (2015) (**SimLex-999**) to evaluate English semantic similarity models. The only available benchmark conceived for evaluating association relations is **WS353-Association**.

Freitas, Barzegar, Sales, Handschuh, and Davis (2016) provides translation for some of those benchmarks in French, however the provided proximity scores are heterogeneous with the original rankings. Scores for French **SimLex-99** are given by a computer semantic model, while **WS353** scores are identical with English data. The latter configuration is problematic since in different languages the translation are not exact mappings between words, and the proximities are subject to the nuanced translation choice. We manually corrected the erroneous translation of word pairs, eliminated and replaced distinct original word-pairs that are translated to the same target word-pairs, or word-pairs to the same words. Scores for the replaced word-pairs are retrieved from the English dataset. The built French benchmark data suffer from the lack of real human judgement data, thus they serve merely as indicators of semantic model performance. The modified benchmarks are made available on GitHub¹⁰.

3.3 fMRI Voxel-Wise Encoding

In this project and many other similar works [TODO, refs], we consider the BOLD signal of a given voxel j as a temporal signal, which is linearly composed by various independent functional sub-signals, which themselves are convolutions of separate neuron activations with a hemodynamic function:

$$\text{BOLD}_{\text{theory},j}(t) = \sum_i \beta_{i,j} \times f_i(t) * \text{hrf}(t), \quad (3.3)$$

where $\beta_{i,j}$ is the linear coefficient of the i -th component, f_i is a function modeling the i -th independent functional activation, and hrf is the hemodynamic function¹¹ (we

¹⁰commit c97583f, <https://github.com/nicolasying/Similarity-Association-Benchmarks>

¹¹We used packaged functions from `nistats`, `nilearn` to implement regressor construction (Abraham et al., 2014).

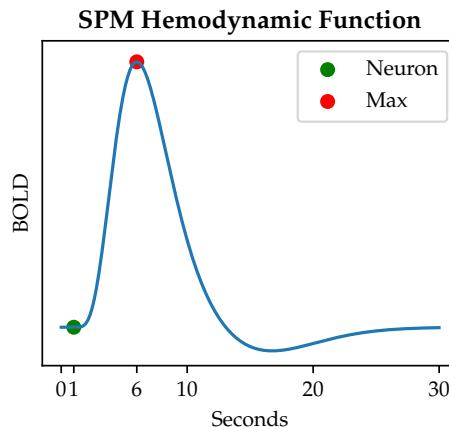


FIGURE 3.3: The shape of the hemodynamic function used in the project. The modeled neural activity instantaneously fires at 1s. The maximum of the hemodynamic function is attained at 6s.

use the model used in SPM at oversampling rate of 10, the hemodynamic function is illustrated in Figure 3.3).

Since different voxels contain neurons of distinct (yet possibly similar) activation profiles towards stimuli, the coefficient associated with each functional component are also different for each voxel. For example, in an auditory comprehension task, the statistical distribution of GLM trained coefficients (β_i in equation 3.3) associated to a voxel which is not implicated in audition to be near 0. Voxels containing neurons primarily associated with low-level auditory functions having non-zero β s for low-level auditory features.

3.3.1 fMRI Textual Stimuli Preparation

For the generation of regression features for fMRI encoding, we perform a lemmatization of the Little Prince story. First the text is parsed with syntactic dependency analysis with spaCy¹² library and each token is attributed with a POS tag. We then used FrenchLefffLemmatizer¹³(Sagot, 2010) library to return verbs to the infinitive form and other words to masculine singular form. POS info from spaCy helps to resolve lemmatisation ambiguity. The pipeline-generated lemma and POS tags are then manually verified and corrected¹⁴.

3.3.2 Regression Feature Generation

The exact transcription of the audiobook is performed by Todorovic and Pallier (2018). With jtrans and Praat, the authors aligned the audio with the text by marking the onset and offset of each pronounced word in the story. To pin down the BOLD signal at a given time, we reconstruct temporal regressor functions by firstly build f_i in equation 3.3, which is essentially a sequence of bumps each occurs at the onset of a word (or content word), then we concatenate the sequence of convoluted D regressors into a design matrix of D columns.

¹²version 2.0.16

¹³commit ba1ef2b. The library is publicly available on GitHub, <https://github.com/ClaudeCoulombe/FrenchLefffLemmatizer>.

¹⁴The pipeline and hand-made modifications are available at <https://github.com/nicolasying/Micipsa-Text-Preprocessing/blob/master/Text%20PreProcessing.ipynb>.

We used four groups of features to reconstruct human listening comprehension processing cerebral activities:

1. RMS (Acoustic Energy), which is the root means square of audio wave amplitude calculated on a sliding window of 10 msec, with Octave¹⁵.
2. WRATE (Word Presence), a binary temporal sequence indicating if a word is being pronounced at a given time.
3. CWRATE (Content Word Presence), a similar binary feature to WRATE, which indicates the presence of a content word, determined by the POS tag (including nouns, verbs, adjectives and adverbs) of the lemmatised text.
4. SIM/ASN/SIG/MIX (Semantic Embedding-Based Features), a multi-dimensional feature set. The feature value is taken from a certain embedding defined in $\mathbb{R}^{|vocabulary| \times n}$ space, where a particular matrix row corresponds to the content word in question.

RMS and WRATE are reported by post-hoc analysis of Todorovic and Pallier (2018) as informative features. Henceforth, we define a *regressor group* the regressors built from a group of features, *regressor class* as the combination of regressors from the regressor group with the same name and the groups of lower feature levels. For example, the *regressor class* SIM contains regressors from *regressor groups* of RMS, WRATE, together with SIM.

For the ease of later design-matrix feature selection, we systematically performed orthonormalization of the convoluted feature sequences to cancel the co-linearity of the regressors. This is implemented with Gram-Schmidt process¹⁶ (“GramSchmidt Process,” 2019), where the orthogonal sequence is defined by the order of regressor classes above, and the semantic embedding based regressors inner-class order is either given by the original semantic model (ASN/SIG/MIX) or by PCA (SIM).

3.3.3 Feature Selection for Specific Corpus

SIM space is constructed by taking the first PCs of a transformed ontological graph adjacency matrix, the explained variance associated to each PC is informative. As *Le Petit Prince* uses limited vocabulary, the semantic space might not be fully exploited as it was factorized with a much larger lexicon. We suspect that there might be some semantic dimensions in the semantic spaces that are not fully exhibited. It is in our interest to simplify the design matrix, leave out uninformative feature columns (those with extremely low variances) to avoid overfitting and accelerate model fitting computation. Therefore we take an investigation of the 9 design matrices (one per fMRI block) by averaging each design matrix’s variance of individual regressors. After orthonormalization, the variance of regressors in higher dimension positions are of a much smaller order than the first regressors especially for PCA-factored semantic spaces. The value of threshold for column selection is determined visually to limit the number of informative regressors under 200.

¹⁵<https://www.gnu.org/software/octave>

¹⁶Gram-Schmidt process is an iterative procedure applied over a set of linearly independent functions. It construct an orthogonal basis by subtracting the projection of a posteriorly positioned column over existing partial orthogonal basis (initially this basis is the first column), so that the residual of the subtracted column is linearly independent. The column residual is added to the orthogonal basis and the procedure continues until the last column is added to the basis.

3.3.4 Ridge Regression with Step-wise Forward Feature Selection, Grid Search and Cross Validation

The fMRI encoding protocol is to find a function projecting our theoretical feature regressors onto BOLD amplitudes. We assume that the target BOLD value is linearly composed by individual regressors, similar in equation 3.3. The coefficients in the equation above is determined by the minimization of the squared difference between the predicted value given by the voxel-model and the recorded BOLD value, on a set of discrete timestamps. This is a typical regression problem tackled in Machine Learning. The computation of the coefficients is named *training* or *fitting* of the model. The performance of a fitted model trained on a dataset is evaluated on the accuracy of its predictions on unseen data, which indicates its *generalized predictive power*.

In each fMRI recording block we have around 300 whole-brain images (refer to Section A.2 for more details), totaling 2937 observations. Researchers (Hua, Xiong, Lowey, Suh, & Dougherty, 2005) found that in GLM the optimal number of uncorrelated informative features is $N - 1$ where N is the number of observation, and \sqrt{N} if features are correlated. Although we have numerically de-correlated the regressors, we nevertheless cannot assume the conceptual independency. Thus 200 regressors might outnumber the recommended feature size. To avoid potential overfitting of regression models, we use Ridge regression to penalize the attribution of large coefficients. Equation 3.4 is the minimization problem posed by Ridge regression, with j fixed for voxel j and N_j the number of features used by voxel j . Strong penalization reduces potential noises by limiting the chance of particular feature columns weighing too much on final prediction, therefore promotes the robustness and generalizability of a fitted model.

$$\min_{\beta_{i,j}} \sum_t \left| \sum_i \beta_{i,j} \times f_i(t) * \text{hrf}(t) - \text{BOLD}_{\text{real},j}(t) \right|^2 + \alpha_j \sum_i \beta_{i,j}^2 \quad (3.4)$$

Ridge regression fitting algorithm requires a hyper-parameter (α) to adjust the severity of large-coefficient penalty. There's no empirically predetermined optimal choice of the value for similar project settings, thus we test a range of candidates by fitting different models and evaluate their performance.

We also assumed the heterogeneity of voxel activation profile towards different functional features. In order to maximize the predictive power of models, we trial multiple combinations of feature columns on each voxel-model. Limited by the computation time, we do not test all the combinations of individual features which could results in an exponential complexity, but use *step-wise forward* feature selection by the order of feature classes (Section 3.3.2).

A major difference of this project from Todorovic and Pallier (2018), Verdier et al. (2018) is the voxel-specific configuration. Our pilot regression experiments partially replicated the original experience (Todorovic & Pallier, 2018) on French data. The results showed that fixing one regularization parameter for all models (including non-semantic models and MIX models) cannot fully exploit the power of the supplied regressors. The fixed α preferentially improves the regression performance for certain models.

By prior experiences, a search range for α is fixed beforehand. We sampled 34 α s from the defined range and up to 32 feature dimension candidates, in hope of including the near-optimal hyper-parameter combination for the regression of each voxel-level model. The list of tested parameters are fixed in each model's config file¹⁷.

In our project, for each voxel of a subject and each combination of α value candidate and feature selection, we adopt the common practice of Cross Validation (CV), where we generate 9 different regression models by training on 8 runs of fMRI recording leaving one out for validation, and test their performance on the left run by computing the coefficient of determination (r^2) by comparing model predicted BOLD values and real observations. We will henceforth name the model validated on fMRI block i *run i*. r^2 measures the proportion of the variance in the BOLD signal that is predictable from the feature regressor data. An r^2 of 1 indicates that the regression predictions perfectly fit the data.

We normalized all feature regressors and voxel-wise fMRI signal sequences for the facility of inter-individual comparison and group-level analysis. To reduce the total computation time, we filtered out unimportant voxels in the images by computing a multi-EPI mask.¹⁸

3.4 Analysis

The Ridge regression pipeline results to

$$|\text{Subject}| \times |\text{Voxels}| \times |\alpha \text{ candidates}| \times |\text{feature selection}| \times |\text{CV}|$$

fitted voxel-models for each semantic space.

3.4.1 Incremental Nested Model Sequence

First we investigate the validity of regression results.

For each individual voxel in 9 cross-validation sessions, $|\alpha \text{ candidates}| \times |\text{feature selection}|$ results are given. For each cross-validated model, we select the highest r^2 among α s and feature selections within feature classes. For example, for voxel j CWRATE feature class result, we take the maximal r^2 score for each voxel j model among all the scores found with RMS, RMS+WRATE and RMS+WRATE+CWRATE features and all tested α s. This selection reduces the number of reported scores to $|\text{Subject}| \times |\text{Voxels}| \times |\text{feature classes}| \times |\text{CV}|$. This reporting approach is proposed due to the overfitting problem of regression models: the addition of extra features does not necessarily translate into a higher performance. If a model overfits (i.e. r^2 declines) with the addition of features, the overfitted nesting-model results are substituted with un-overfitted nested-model ones so that on whole-brain maps the best voxel-models are always visualized. For example, if a voxel's r^2 performances with different regressor classes are ranked as follows: CWRATE > SIM > RMS = WRATE, then in whole-brain visualization and model comparative analyses, WRATE r^2 is used for RMS and WRATE,

¹⁷For example, ASN is configured as <https://github.com/nicolasying/Micipsa/blob/master/models/fr/rms-wrate-cwrate ASN200/config.json>. Please refer to Section A.3 for tested α value, feature selections.

¹⁸The `nilearn.masking.compute_multi_epi_mask` uses the mask-finding algorithms to extract masks for each session of subject, and then keeps only the main connected component of a given fraction of the intersection of all the masks.

and CWRATE r2 for both CWRATE and SIM. Thus the contrast of SIM against CWRATE is zero rather than a negative number.

For each feature class, the subject-wise result whole-brain map and the group-wise map are visualized by averaging across |CV| and |Subject| \times |CV| voxel- and feature-class-specific results. With each additional feature class starting from WRATE, the improvement of r2 scores are also plotted. With the downward-inclusive best model selection, only non-negative contrasts are reported.

$$F = \frac{\frac{RSS_{\text{restricted}} - RSS_{\text{full}}}{p_{\text{full}} - p_{\text{restricted}}}}{\frac{RSS_{\text{full}}}{n - p_{\text{full}}}}, \quad (3.5)$$

where p is number of features, n is number of samples.

The statistical significance of improvement is computed by Wald F-test (Equation 3.5 on model validation scores. The Wald F-test compares the residual sum of squares (RSS) of a restricted model and a full model nesting the former one, with the null hypothesis suggesting that the full model does not provide a significantly better data fit than the restricted one. The Wald test penalizes large feature set, and takes the number of observations into account, thus is more restrict than tests comparing r2 scores. The RSS is computed from r2 given that the data are centered and normalized (Equation 3.6 for voxel j).

$$\begin{aligned} RSS_j &= \sum_{t=0}^n (BOLD_{\text{real},j}(t) - BOLD_{\text{predict},j}(t))^2 \\ r2_j &= 1 - \frac{\sum_{t=0}^n (BOLD_{\text{real},j}(t) - BOLD_{\text{predict},j}(t))^2}{\sum_{t=0}^n (BOLD_{\text{real},j}(t) - BOLD_{\text{average},j})^2} \\ &= 1 - \sum_{t=0}^n (BOLD_{\text{real},j}(t) - BOLD_{\text{predict},j}(t))^2 \\ &= 1 - RSS_j \end{aligned} \quad (3.6)$$

For each addition of *feature group*, we perform a Wald F-test within each cross-validation session for each voxel. The full and restricted model scores are selected among $|\alpha \text{ candidates}| \times |\text{feature selection}|$ within the corresponding *feature class*. The number of features of each model are determined by feature selection, and the number of samples is the fMRI image number of the cross-validation session.

At individual analysis level, for each contrast of each individual voxel, |CV| F-tests are computed. For the final significance visualization, we compute the geometric mean of p-values over |CV| runs. We then plotted the statistical map by thresholding the p-value by uncorrected 0.05, uncorrected 0.001 and Bonferroni multi-comparison corrected 0.05. For group analysis, the geometric mean is computed over |subject| \times |CV| observations.

To pin down cortical regions well modeled by a particular class of model, we select the best 0.1% and 1% voxels and report voxel-clusters larger than 1500 mm^3 (47 voxels). For r2 difference maps, smaller regions are permitted (500 mm^3 (16 voxels)), the lower bound of voxel-wise Wilcoxon statistic significance of the clusters are also

reported. With F-test results, we report voxels surviving three-levels of significance thresholds.

Without prior hypothesis on semantic embeddings, the regression pipeline is expected to recover at least auditory cortical areas for RMS feature regression by plotting the whole-brain map of r^2 . The additional features' and contrast maps' validity are backed by the validity of RMS activation regularities.

3.4.2 Embedding Contrasts

The key comparison of regression results is the contrast between SIM and ASN semantic models. The comparison computational procedure is inspired by the non-nested model comparison pipeline (Merkle, You, & Preacher, 2016). Our method is divided into two steps: first we verify the structural difference between the design matrices given by each model, secondly we compare the model's regression results if the design matrices are found nonequivalent. The pipeline is detailed in Section A.4.1.

For voxel-model regression result contrasts, the r^2 s follow distributions described by $F(k - 1, n - k)$, where k is the number of features and n is the number of observation. Since across semantic models, the feature dimensions are heterogeneous, r^2 scores have distinct score distributions. Therefore we adopted the nonparametric Wilcoxon signed-rank test to test the significance of r^2 -differences between semantic models.

For voxel-level group contrasts, we take two paired groups of r^2 scores, each composed by $|\text{Subject}| \times |\text{CV}|$ observations. The Wilcoxon test yields a W statistic and a p-value for each voxel. For individual contrasts, since $|\text{CV}| < 20$, the Wilcoxon test is tested with T statistic. For a group size of 9 observations, T critical values for two-tailed alternative hypothesis are 8, 5, 3, 1 for alpha (statistic power) < 0.1, 0.05, 0.02, 0.01.

In result

3.4.3 ROI-level Analysis

With Region-of-Interest (ROI) analysis, we aim to filter out inter-subject variances and find relatively stable loci of each semantic network. We replicated Patterson et al. (2007)'s peak selection method to select the reported peaks in temporal region from the relevant literatures reporting the activation peaks of semantic tasks (Binder et al., 2000; Bright, Moss, & Tyler, 2004; Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003; Davis & Johnsrude, 2003; J. T. Devlin et al., 2000; Ferstl, Rinck, & von Cramon, 2005; M. L. Gorno-Tempini & Price, 2001; Maria Luisa Gorno-Tempini et al., 1998; Grossman et al., 2003; Mummery Catherine J., Patterson Karalyn, Hodges John R., & Wise Richard J. S., 1996; Mummery, Shallice, & Price, 1999; K. Nakamura et al., 2000; Katsuki Nakamura et al., 2001; Noppeney & Price, 2002; Papathanassiou et al., 2000; Price et al., 2005; Rogers et al., 2006; Scott, Blank, Rosen, & Wise, 2000; Scott, Rosen, Lang, & Wise, 2006; Simons, Koutstaal, Prince, Wagner, & Schacter, 2003; Small, Jones-Gotman, Zatorre, Petrides, & Evans, 1997; Tranel, Grabowski, Lyon, & Damasio, 2005; Tsukiura, Mochizuki-Kawai, & Fujii, 2006; Vuilleumier, Henson, Driver, & Dolan, 2002), and obtained the list of ROI by constructing a 7-mm diameter sphere around the peaks.

We completed the ROI list by intersecting the original ROI mask with gray-matter mask, and added classic language-related brain anatomical areas including Superior

Temporal Gyrus (STG), Inferior Frontal Gyrus (IFG), IFG pars opercularis (IFGoper), IFG pars orbitalis, IFG pars triangularis, laLFissure [TODO], Temporal Lobe (TL), Temporal Pole (TP), posterior Superior Temporal Sulcus, Temporoparietal Junction (TPJ), anterior TL, Putamen, Middle TG, left Premotor Cortex,(Pallier, Devauchelle, & Dehaene, 2011). The ROI centroids are reported in Section A.4.2.

ROI Statistical Test

We computed the ROI-average r^2 , and used the same Wilcoxon signed-rank test as voxel-wise analysis.

Chapter 4

Results

4.1 Semantic Embeddings

4.1.1 Validation on English Data

For SIM space, we used the English WordNetEmbedding trained on the first 15,000 frequent words and benchmark dataset vocabulary. We kept first 511 PCs by following the method in Section 3.2.1, which is comparable to the best dimensionality (850) reported by the original work (Saedi et al., 2018) for a WordNetEmbedding keeping all semantic relations.

The intersection of SIM space and the Common Crawl vocabulary used in MIX space resulted to 8157 words.

The linear regression model mapping SIM to MIX produced a r2 score of 0.1662.

Figure 4.1 shows the PCA-factored explained variances of the 4 resulting semantic space PCs. Table 4.1 shows the semantic ranking task evaluations. Conformably to our hypotheses, the untouched MIX space is indeed a mixture of *similarity* and *association* information. SIM, which is constructed with specifically picked semantic relations from WordNet, is purely *similarity* with statistically negligible *association* scores. Though we have not completely purged *similarity* information from ASN and *association* from SIG, each resulting semantic space has significantly reduced the score in its irrelevant semantic axis. In addition, a clear dominance of *association* semantic signal is present in ASN and *similarity* in SIG.

4.1.2 Application on French data

Provided with the methodological success of English data, we applied the same algorithm against French data.

For SIM space, we used the French WOLFEembedding with POS tag trained on all the available vocabulary. We kept first 634 PCs.

After rule-based and manual matching, the intersection of SIM space and the MIX space vocabulary resulted to 24519 distinct lemma with POS tags.

The linear regression model mapping SIM to MIX produced a r2 score of 0.0776, which is lower than the English score, indicating a smaller informational overlap between the two embedding models.

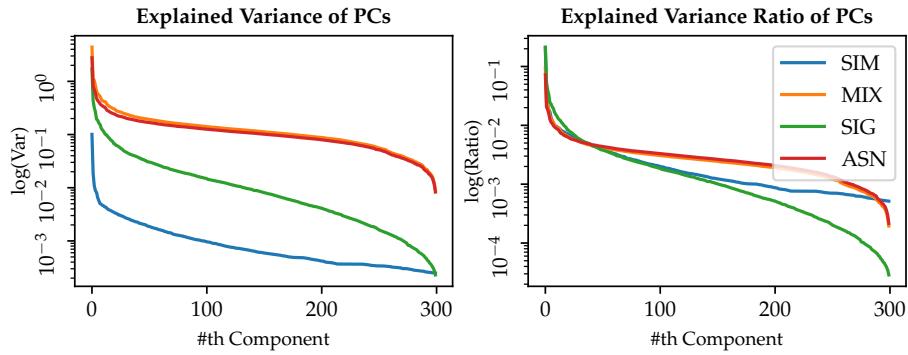


FIGURE 4.1: PCA of 4 semantic spaces of the English 8157 vocabulary. **Left panel:** While the variance of SIM is systematically smaller than other three spaces, its projection (SIG) has larger variances. The suppression of SIG from MIX has little impact on MIX’s variance. **Right panel:** SIG and SIM have a denser variance concentrated on first PCs, while ASN and MIX have more homogeneous variance distributions.

English Semantic Space Semantic Ranking Task Results

Semantic Space	Vocabulary Size	Dimension	r	SimLex-999	WS353-ASN
SIM	15K	511	Pearson	.5060	.0279 ¹
			Spearman	.4989	.0193 ²
MIX	2.2M	300	Pearson	.3946	.6091
			Spearman	.3752	.5709
ASN	8157	300	Pearson	.1953	.5633
			Spearman	.2133	.5918
SIG		Out of Vocabulary	Pearson	.4929	.2091
			Spearman	.4994	.1678
				.002	.024
Baseline ³	13k	850	Pearson	.50	.32
			Spearman	.52	.33

Scores marked in bold have a p-value larger than 0.05.

¹ p-value=0.6626

² p-value=0.7629

³ Baseline is reported by Saedi, Branco, António Rodrigues, and Silva (2018). The 13k words are selected cue words in psycholinguistic experiments. They show the best performance among all tested models.

TABLE 4.1: With a different semantic relation selection, SIM achieves almost the same performance as the baseline in *similarity* benchmark, while it cancels out the *association* score. MIX space performs well in both task-sets, with a slight preference for *association*, consistent with (Lapesa, Evert, & Schulte im Walde, 2014)’s conclusion. ASN has comparable scores in *association* with MIX, but still have a non-zero score in *similarity*. The projected SIG space compared with SIM has similar scores in *similarity* and a much lower score in *association*.

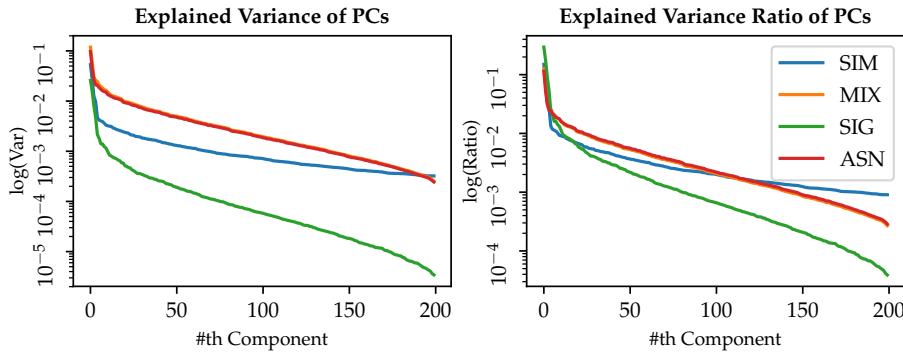


FIGURE 4.2: PCA of 4 semantic spaces of the French 24519 vocabulary. **Left panel:** Due to the poor linear correlation found between SIM and MIX, the variance of SIG is systematically smaller than the other three spaces, the original space SIM has larger variances. The suppression of SIG from MIX has also little impact on the model’s variance. **Right panel:** SIG has a denser variance concentrated on first PCs, while the other three spaces have more homogeneous variance distributions.

Figure 4.2 shows a similar PCA explained variance distribution. Though no sound evidence supporting the validity of the French benchmarks, we tested the resulting semantic spaces using the same tasks against our indicative gold-standard data (Section 3.2.3). The unmodified MIX space has a much lower score in *similarity* and *association* compared to the English MIX, setting a weak baseline for embedding space comparison. Again, SIM achieves high scores in *similarity*, and negligible scores in *association*. However, the results (Table 4.2) puts the validity of ASN space into question: viewed by Pearson’s r it seems to contain none of *similarity* and *association* information, judged by Spearman’s r , both axes’ information are present in the space. All p-values reported for ASN are close to the significance threshold. However, *association* scores of ASN are nevertheless higher than *similarity* scores. Despite that SIG represents only a very small portion of variances of the MIX embedding as indicated by a low r^2 score, SIG has comparable *similarity* scores as SIM, and the purity against *association* is even more remarked.

To further control the quality of the resulting spaces, particularly that of ASN due to debatable scores, we visualized the French semantic spaces with an embedding projector¹ to visualize several exemplar lexicon units and its vectorial neighbors. Based on this analysis (examples in Section B.1.1), we are convinced that French ASN has a predominant *association* preference.

4.2 Computational Analysis of Ridge Regression

4.2.1 Regressor Generation

Vocabulary Coverage

Each word (lemma) in the narrated story used in fMRI experience is associated with its RMS acoustic feature temporal evolution and its semantic values in different spaces. However some of the content words are not all available in our obtained

¹Published as a TensorFlow component, available at <https://projector.tensorflow.org/>. The entries in the embedding space is presented by a sphere positioned in a 3D space, of which the coordinates are by default calculated with the first 3 PCs.

French Semantic Space Semantic Ranking Task Results

Semantic Space	Vocabulary Size	Dimension	r	SimLex-999	WS353-ASN
SIM	56665	634	Pearson	.3291	.1039
			p-value	0	.1061
			Spearman	.2812	.0511
			p-value	0	.4273
Out of Vocabulary			.048	.04	
MIX			Pearson	.0940	.1520
			p-value	.0047	.0197
			Spearman	.1449	.2078
			p-value	0	.0014
ASN	24519	200	Pearson	.0629	.1116
			p-value	.0590	.0879
			Spearman	.0771	.1566
			p-value	.0206	.0162
SIG			Pearson	.2541	-.0044
			p-value	0	.9458
			Spearman	.3121	-.0078
			p-value	0	.9050
Out of Vocabulary			.0797	.0711	

Scores marked in bold have a p-value larger than 0.05.

TABLE 4.2: The results are consistent with English semantic spaces, despite the poor quality of French benchmark datasets. SIM has high performance in *similarity* and negligible *association* scores. The relatively poor de-correlation between SIM and MIX resulted a ASN still containing abundant *similarity* information. SIG however, cancels out completely *association* information even compared with SIM while retained *similarity* signals.

The Little Prince Vocabulary Coverage

		# Instances in fMRI Recording Session								
		R1	R2	R3	R4	R5	R6	R7	R8	R9
Story	T	725	812	860	762	732	902	819	712	802
	V	348	360	411	329	292	367	302	328	343
SIM	TM	36	30	32	27	30	33	24	30	27
	%	4.97	3.69	3.72	3.54	4.10	3.66	2.93	4.21	3.37
56665	VM	26	16	22	20	16	19	16	19	16
	%	7.47	4.44	5.35	6.08	5.48	5.18	5.30	5.79	4.66
ASN	TM	48	47	38	37	48	60	35	37	41
	%	6.62	5.79	4.42	4.86	6.56	6.65	4.27	5.20	5.11
/SIG	VM	30	26	26	25	26	32	20	22	25
	%	8.62	7.22	6.33	7.60	8.90	8.72	6.62	6.71	7.29

TABLE 4.3: Only content words are taken into consideration. T: Token, V: Distinct Lexicon Unit, M: Miss

spaces (Table 4.3 and B.3). When generating regressors, the semantic vectors are set to zero for out-of-vocabulary lemmas.

Corpus-Targeted Semantic Feature Selection

The PCA dimension cutting methods presented in Section 3.2.1 produced 634 feature dimensions for French SIM. After having generated regressors with word onset timestamps and semantic representation vectors, the average variance cross 9 groups of 103 resulting regressors are computed and visualized in Figure 4.3. We selected the threshold of 10^{-5} , which resulted 100 informative regressors for SIM space².

4.2.2 Choice of α and Effective Feature Dimensionality

For each of four semantic models, we generated design matrices for each fMRI session with 103 or 203 features (including 3 non semantic embedding features).

Analysis in Section B.3.1 suggests that our research space for α and feature-dimension parameters are complete: the distribution of voxel-configurations are bounded by our search space (see Figure 4.4 for the MIX distribution of subject 1 for an example, session-wise distribution plotting for all models are available online [TODO, add url]).

Section B.3.1 also illustrates the interaction between α values and voxel regression performances, it confirms that a uniform α setting penalizes certain cognitive models thus creating an important bias for model selection (see Section 3.3.4), supporting our choice of voxel-wise α configuration.

²The selected dimensions are 1 – 85, 87 – 94, 96, 97, 99, 100, 103, 117, 131.

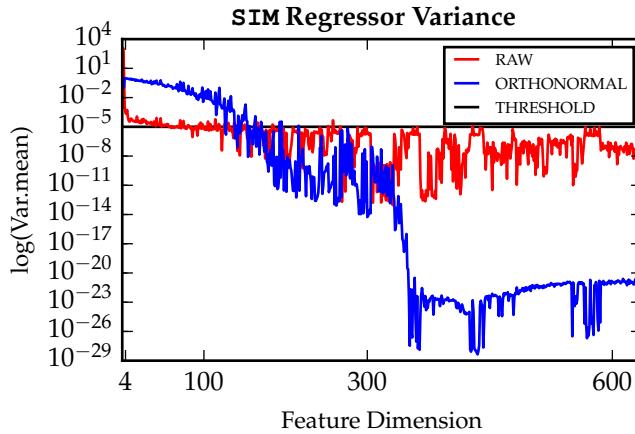


FIGURE 4.3: The average of 9 fMRI run semantic regressor variances. There are 3 non-embedding regressors, and 634 SIM-based regressors. RAW stands for regressor values directly after hemodynamic convolution, ORTHONORMAL stands for de-linearized regressors after Gram-Schmidt process. THRESHOLD for regressor selection is fixed at 10^{-5} . The RAW regressors' variance declines dramatically after first few SIM regressors ($\text{dim} > 3$), and stays relatively stable for later dimensions. This observed trend corresponds well to the eigenvalue evolution of SIM space (Figures 3.1 and 4.2). ORTHONORMAL regressors's variance declines more slowly, and has a noised plateau around dimension $100 - 300$. Later regressors suffer more significantly in variance (smaller than 10^{-23} , approaching the computation precision limit of Python floats) and retained almost no information for the second half PCs. The threshold is cut around the upper bound of the ORTHONORMAL variance plateau noise, so a continuous regressor set could be included in the final design matrix without surpassing the dimensionality limit (of 200 which is the dimensionality of the used DepGlove embedding).

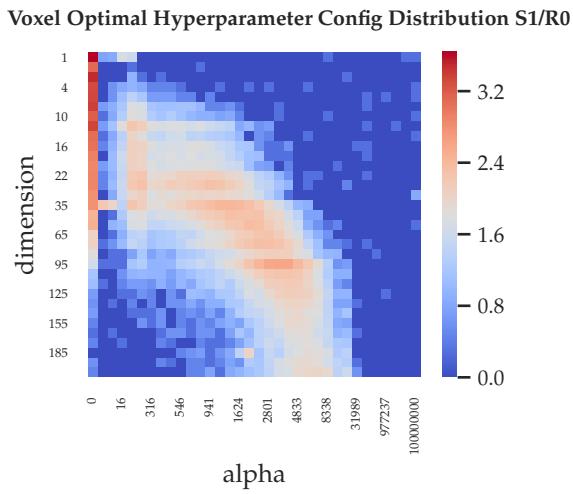


FIGURE 4.4: After averaging MIX best r^2 s across 9 runs of one same subject, best hyper-parameter configuration appears to be regularly distributed in the search space. A large proportion of voxels are best modeled with no Ridge regularization (especially for voxels using < 4 features). Voxel-models requiring for higher-dimensional features are associated with larger α values. A diagonal trend is found bounding α . $\alpha > 10^{4.5}$ (31989) rarely achieves best predictive performances, suggesting that the α search space is complete for the subject.

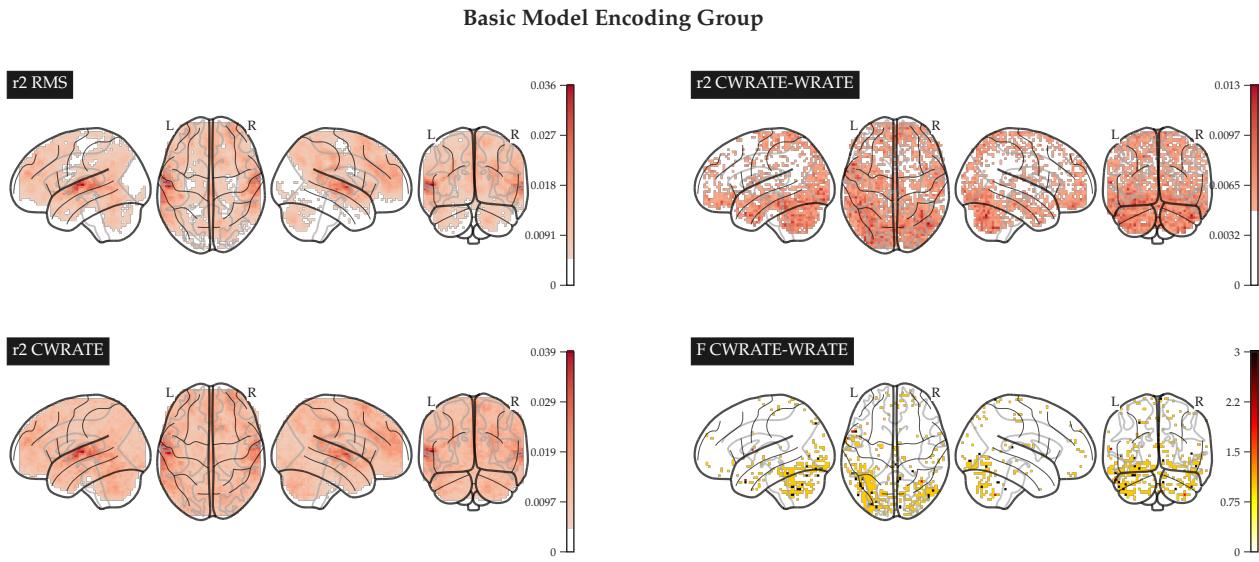


FIGURE 4.5: **Left panels** are plots of voxel-wise r^2 scores for three *classes* of non semantic-embedding regressors. With RMS, WRATE and CWRATE, consistent model performances are found for bilateral primary auditory cortices, with a slight preference for left hemisphere (Table 4.4). **Right upper and mid panels** are r^2 difference maps between two regressor classes. WRATE does not improve any voxel performance, and CWRATE improvements are mainly located in bilateral TP, MTG, ITG, frontopolar PFC and cerebellum near Fusiform cortex (Table 4.5). **Right lower panel** is the F-test contrasting RMS+WRATE+CWRATE and RMS+WRATE. 3 levels of significance 1, 2, 3 are shown on the whole-brain map, corresponding respectively to p-values of uncorrected 0.05, 0.001 and voxel-wise corrected 0.05. The locations of the significant voxels are reported in Table B.7.

4.3 Cognitive Analysis of fMRI Encoding

For convenience, the voxel-model predictive performance is considered as voxel activations in response to different feature sets.

4.3.1 Non Semantic-Embedding Models

Figure 4.5 illustrates the group-wise results for non-semantic-embedding feature classes. Subject-wise results are available online³. Additionally, Figure 4.6 shows the exact impact on addition of WRATE and CWRATE *feature groups* to the design matrix, without substituting worse nesting-model results with better nested-model results. Subject-wise results are available online⁴, the overfitting with feature additions are more pronounced.

RMS preferentially models primary auditory cortical activities in Broadmann Area (BA) 41, together with other two clusters located in right mid cingulum and right middle frontal cortex (Table 4.4). The addition of WRATE does not bring any impact (Figure 4.6 left panel), possibly due to the high co-linearity with RMS by definition, thus the orthonormalized feature contains only uninformative noises despite a relatively important variance (0.96 after orthonormalization, CWRATE has only 0.10 for

³http://bit.ly/micipsa_base_wholebrain

⁴http://bit.ly/micipsa_regression_histogram

RMS/CWRATE/SIM/ASN Best Modeled Voxel Clusters									
Position	BA	Functional Label	Feature Class	x	y	z	#Voxel	r2 Peak	r2 Min
Top .1%									
Temporal Sup L	41	Prim Auditory	RMS	-59	-13	6	56	.0362	.0240
			CWRATE	-61	-12	4	61	.0387	.0259
			SIM	-60	-13	6	53	.0412	.0291
			SIG	-61	-11	4	87	.0430	.0301
			ASN	-60	-13	6	53	.0388	.0293
Top 1%									
Temporal Sup L	41	Prim Auditory	RMS	-60	-12	4	410	.0362	.0150
			CWRATE	-60	-11	4	417	.0387	.0178
			SIM	-60	-11	3	522	.0413	.0214
			SIG	-61	-17	4	527	.0430	.0301
			ASN	-60	-19	4	515	.0388	.0218
Cingulum Mid R	23	-	RMS	0	-24	29	69	.0189	.0150
			CWRATE	1	-23	29	55	.0202	.2587
			SIM	2	-24	44	37 ¹	.0243	.0214
			SIG	2	-25	45	55	.0250	.0301
			ASN	-2					
Frontal Mid R	10	-	RMS	29	56	20	103	.0180	.0150
			CWRATE	28	60	19	55	.0221	.0178
			SIM	29	59	19	92	.0256	.0214
			SIG	26	57	20	86	.0252	.0301
			ASN	30	59	19	74	.0259	.0218
Temporal Sup R	41	Prim Auditory	RMS	61	-13	2	337	.0328	.0150
			CWRATE	62	-11	3	198	.0342	.0178
			SIM	62	-9	1	264	.0366	.0214
			SIG	61	-11	3	262	.0402	.0301
			ASN	62	-13	3	266	.0370	.0218

Coordinates are reported in Montreal Neurological Institute (MNI) standardized spaces.

¹ Cluster size smaller than 47 voxels, 1500 mm³.

² Cluster not found.

TABLE 4.4: Across four regressor classes, the best models voxels are consistently found in four cortical areas: bilateral pSTG (BA41, primary auditory), right Mid Cingulum and right aPFC (BA10). The most severe voxel score selection leads to left primary cortex (BA41) activation. With the addition of semantic features including CWRATE, SIM and ASN, left pSTG cluster grows while right MCingulum shrinks. CWRATE penalizes voxels in right BA10 and BA41, while semantic regressors re-improves the scores. With ASN, a significant left pSTG cluster centroid posterior shift is observed. No other contrasting differences are found for SIM and ASN.

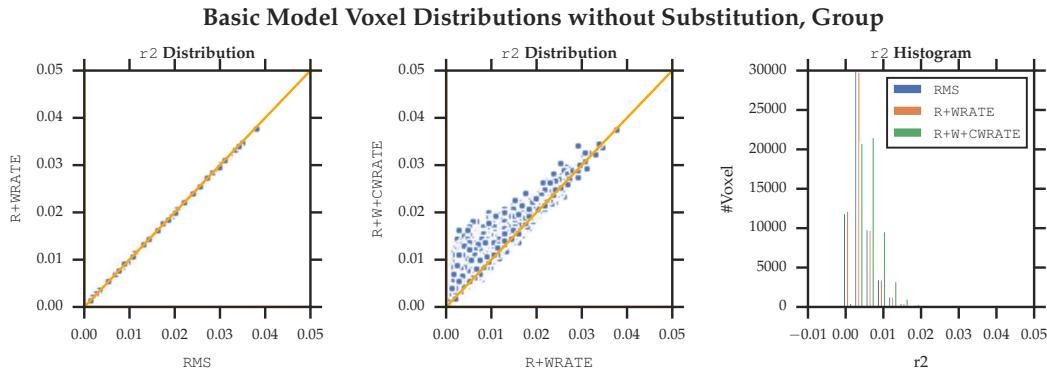


FIGURE 4.6: The voxel scores are averaged over all cross-validations of all subjects. **Left panel** shows that the addition of WRATE does not improve any voxel's model performance. **Mid panel** suggests that CWRATE slightly overfits a small portion of voxels, the improvement for most voxels are minute. **Right panel**: However for originally randomly-modeled voxels (x-axis from 0–0.01), CWRATE does bring significant improvements. The effect of model substitution is not pronounced in group average. However, subject-wise results are significantly higher (up to 0.2) and the score variability and overfitting are more remarkable.

CWRATE Best Improved Voxel Clusters

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2 Peak	$-\log_{10}$ p-value
Top 2%							>.0067	>3.66
Temporal Pole Mid L	38	-	-53	11	-33	22	.0118	4.18
Temporal Inf L	37	Fusiform	-47	-43	-24	90	.0114	4.18
Rectus L	11	-	-5	46	-26	16	.0110	4.00
Cerebellum Crus2 R	37	Fusiform	45	-69	-38	89	.0130	4.35

TABLE 4.5: The most severe voxel score selection of RMS leads to left primary cortex (BA41) activation. Also well modeled voxels are distributed in more extensive areas of bilateral BA41 and right BA23 and BA10. With the addition of CWRATE features, voxel performances are systematically improved. With CWRATE, no other clusters appear in the thresholded voxel set. Left BA41 has a higher concentration of best modeled voxels, while right BA41 and right mid cingulum degrade in voxel score ranking. Right BA10 also improves in ranking.

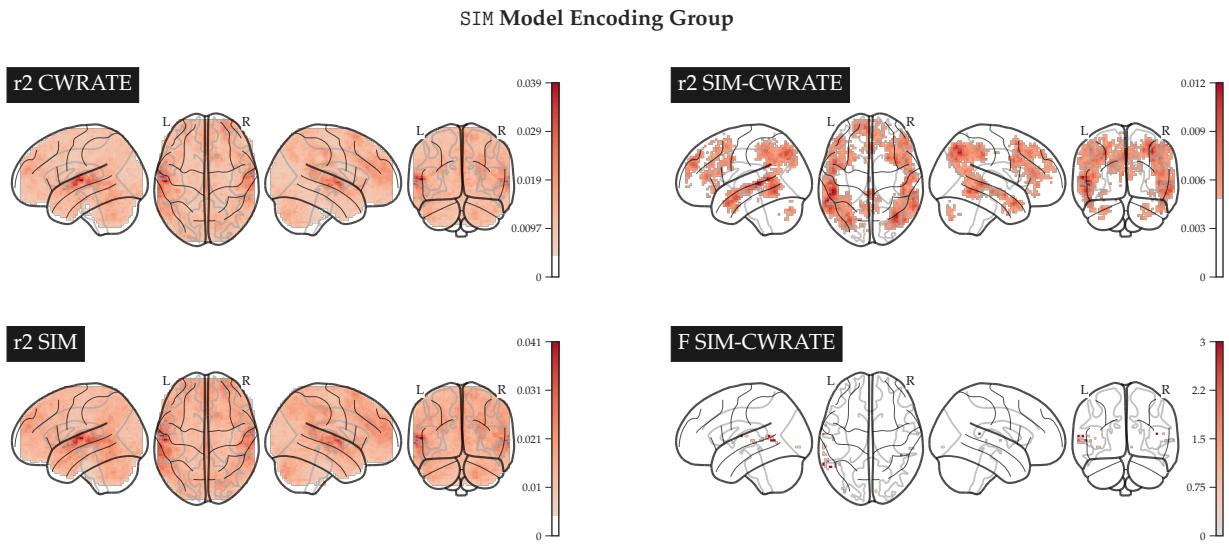


FIGURE 4.7: **Left panels:** The global activation pattern is unchanged with the feature addition. Best modeled zones are bilateral primary auditory cortices. **Right upper panel** shows that SIM better models bilateral MTG, sup Parietal, Angular Gyrus (part of Wernicke’s area), supramarginal gyrus and prefrontal areas (Table 4.6). F-test in **right lower panel** reports significant voxels in left pMTG BA21, 39, right pSTG BA22 and left Heschl BA4 (Table B.7). Subject-wise results are available online at http://bit.ly/micipsa_sim_wholebrain.

reference). With CWRATE, an extensive range of voxels distributed in the whole brain received better performance (Figure 4.6 middle panel).

A left-hemisphere preference for textual listening comprehension is suggested: left primary auditory cortex (BA41) is better modeled than in right hemisphere with RMS and CWRATE. On adding CWRATE, the imbalance between left and right BA41 is enlarged. Table 4.5 reports 4 clusters containing more than 16 voxels that are improved ($W=136$, $\Delta r^2 > 0.0067$, $p\text{-value} < 10^{-3.66}$ uncorrected) in left BA37 (MTP), bilateral BA37 (Fusiform Gyrus) and left Rectus. The Wald F-test on CWRATE contrast reports isolated voxels surviving voxel-wise multi-comparison significance test (Table B.7, $p\text{-value} < 0.05$ voxel-wise multi-comparison corrected). The voxels are mainly located in bilateral BA37, left BA18, 19.

4.3.2 Similarity Nested Model

We added with upon non semantic-embedding models SIM features to construct *similarity* semantic models. While the whole-brain activation pattern stays globally unchanged (Figure 4.7 for group-wise average), in SIM voxel-models, left primary cortex are better ranked than in BASE model, while right mid cingulum models degrade (Table 4.4). SIM enlarges the performance superiority of left STG over right STG, indicating a left preference for textual semantic *similarity* processing. The shrinkage of Mid Cingulum’s proportion in top 1% voxel models might imply that it has a limited participating in *similarity* processing. The r^2 distribution analysis (Figure 4.8 left) shows that in group-average SIM is informative for most of the voxel-models and none of voxels is overfitted by this addition. Table 4.6 reports the most improved voxel clusters by SIM to be located in bilateral MTG, left Sup Parietal and

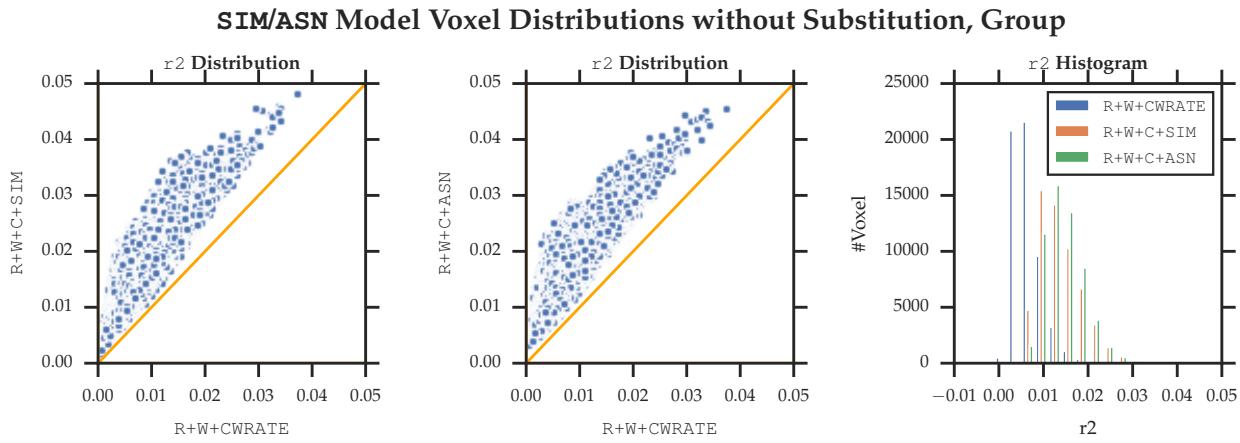


FIGURE 4.8: The group average of semantic embedding models make both important contributions for voxel-modeling (**left** and **mid** panels). **Right** panel shows that ASN model scores are distributed around a higher average (0.014) than SIM (0.01). Subject-wise results are available online at http://bit.ly/micipsa_regression_histogram.

SIM Best Improved Voxel Clusters

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2 Peak	$-\log_{10}$ p-value
Top .5%								
Temporal Mid L	21	-	-51	-34	1	86	.0119	4.35
Parietal Sup L	7	-	-27	-72	44	17	.0099	4.35
Angular R	39	-	35	-65	44	49	.0114	4.35
Temporal Mid R	21	-	57	-36	-0	17	.0098	4.35

TABLE 4.6: We thresholded Wilcoxon signed-rank test's significance at $10^{-4.35}$ as a clean cut is found in p-value histogram, which leads to a selection of top .5% important voxel-model improvements. The largest and most improved voxel-cluster is found in left BA21, then in right angular gyrus which is part of Wernicke's area. A more lateral and smaller-cluster improvement is found in right MTG.

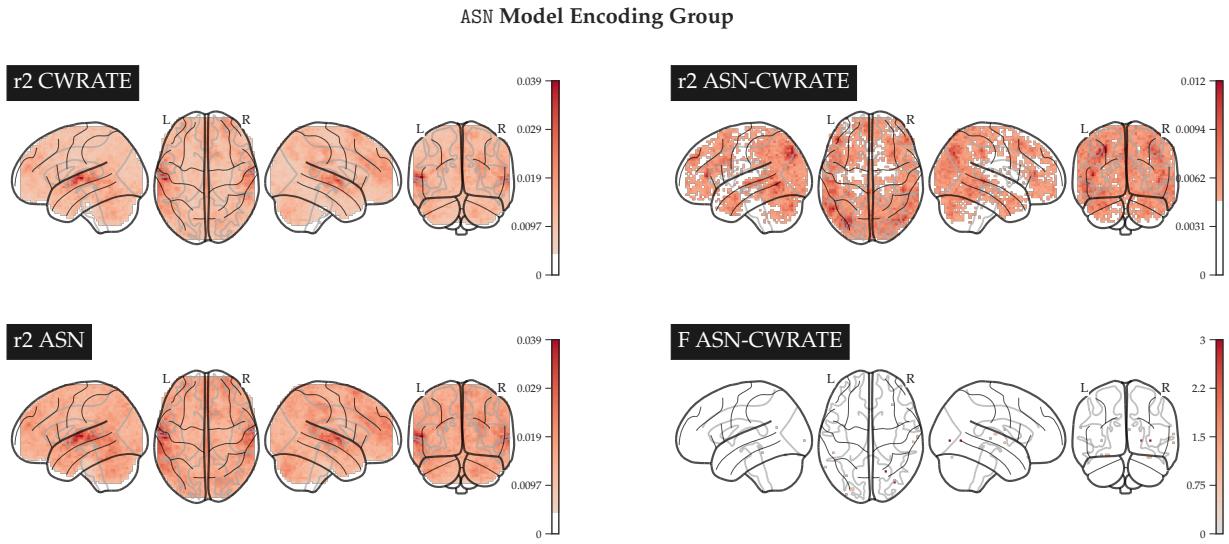


FIGURE 4.9: **Left panels:** Consistent with former results, the addition of ASN on BASE regressors does not change the bilateral primary auditory cortices' dominance. Improvements of voxel-models are distributed in an extensive part of all lobes (**right upper panel**). The most improved voxels are located in bilateral MTG, IFGtri, mid occipital cortex, angular gyrus, sup/mid frontal cortex, mid cingulum (Table B.6). F-test in **right lower panel** reports significant voxels in right lingual BA19 and mid occipital area BA18 (Table B.7). Subject-wise results are available online at http://bit.ly/micipsa_asn_wholebrain.

right Angular Cortex ($W=210$, $\Delta r^2 > 0.0079$, $p < 10^{-4.35}$ uncorrected). Left MTG improvements are more extensive and more important than right MTG. F-test results shows that SIM significantly improves isolated voxels (Table B.7, $p < 0.05$ voxel-wise multi-comparison corrected) in left pMTG BA21, 39, right pSTG BA22.

Similarity Nested Model with SIG

SIG contrasts with CWRATE class models gave similar contrast maps as SIM: bilateral IFGtri, MTG and ITG, SPG and AG are found. With SIG the left ITGtri and left ITG improvements are more drastic. A more detailed presentation of SIG results is available in Section B.4.2.

4.3.3 Association Nested Model

On adding ASN features on BASE features, the bilateral auditory cortices dominance is consistently kept (Figure 4.7 for group-wise average). The contribution brought by ASN on top 4 voxel-clusters initially found by RMS model is similar to SIM (Table 4.4): improvement in bilateral Prim Auditory with a slight left preference, shrinkage of right MCingulum and slight improvement of frontopolar PFC. These clusters therefore do not show an observable preference for SIM and ASN.

ASN brings voxel-model performance boost in an extensive cortical regions. The Wilcoxon test shows near-significant performance improvements ($W=190$, $\Delta r^2 > 0.0065$, $p\text{-value} < 10^{-4.18}$ uncorrected) in left BA39 (visual), right angular gyrus (associated with aphasia), right BA21 MTG, left BA37 fusiform gyrus (Table B.6). F-test results

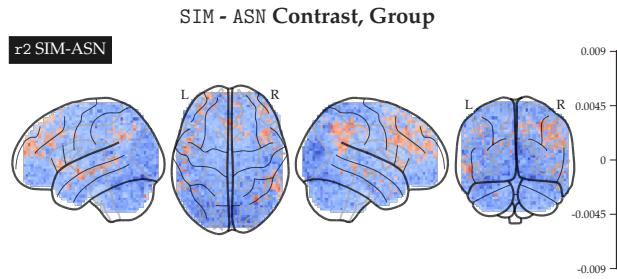


FIGURE 4.10: The differences of best voxel-model r^2 s are plotted. SIM preference is found in left BA10 Superior PFC, left anterior cingulate cortex, left STS, left medial PFC, right IPL, right STG, MTG (Table 4.7). ASN preferences are found in bilateral BA18, right BA20, left BA7, 19, 37 (visual association, primary visual, parahippocampal). Subject-wise results are available online at http://bit.ly/micipsa_sim_asn_contrast

shows that ASN significantly improves isolated voxels (Table B.7, $p\text{-value} < 0.05$ voxel-wise multi-comparison corrected) in right lingual gyrus BA19 and BA18 (visual association).

4.3.4 Similarity/Association Contrast

With SIM

Section B.2 suggests that first feature dimensions of SIM can be partially recovered by ASN model. Therefore, ASN might also be able to model voxels using less than 5 features from SIM, the result might thus lack low-level SIM/ASN contrast. As the first 4 dimensions of SIM encodes primarily POS information (Section B.1.1), we performed ad-hoc regressions on SIM space but uses only lemmas from a certain grammatical category to identify possible impacted regions (upcoming).

The found results are consistent with the conjectures above: ASN scores are higher than SIM in average (Figure 4.8 right), most of voxels respond better to ASN models (Figure 4.10). As the Wilcoxon test shows ($W > 6945$, $\Delta r^2 > 0.0068$, $p\text{-value} < 0.05$ voxel-wise multi-comparison corrected), only two significant clusters are found for SIM in left superior frontal cortex and left anterior cingulum cortex (both associated with control/decision-related cognitions) (Table 4.7) and 17 are found for ASN (Table 4.8) in bilateral visual association areas (BA18), primary visual areas (BA17), ventroposterior temporal areas (fusiform, hippocampus and parahippocampus).

The reported clusters for SIM are composed of 4 to 5 voxels. In our ROI analysis, ROIs larger than 26 voxels are used, thus none of the ROI revealed significance for SIM. As ASN has an overall dominance for almost all brain regions, small ROIs located in left middle/posterior STG and large anatomical structures including IPL and TL all revealed their preference for ASN model.

With SIG

In general SIG [TODO]

SIM-ASN Voxel Contrast, Preference for SIM

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2 Peak	$-\log_{10}$ p-value	Cluster ID
Frontal Sup L	10	-	-27	59	25	4	.0073	6.75	1
Cingulum Ant L	32	-	-8	34	25	5	.0069	6.15	2
Temporal Sup L	22	-	-55	-1	-7	3	.0062	5.35	3
Frontal Mid L	10	-	-36	49	15	4	.0061	5.29	4
Cingulum Ant L	32	-	-2	30	31	2	.0061	5.24	5
Parietal Inf R	40	-	46	-49	44	10	.0059	5.09	6
Angular R	39	-	33	-64	47	2	.0057	4.75	7
Precentral R	8	-	39	8	47	2	.0057	4.73	8
Caudate R	48	Caudate	17	21	3	2	.0052	4.20	9
Frontal Sup R	10	-	27	65	9	2	.0051	4.09	10
Temporal Sup R	39	-	65	-55	22	4	.0050	4.05	11
Frontal Mid L	10	-	-27	40	31	2	.0047	3.68	12
Temporal Mid R	22	-	49	-23	-7	2	.0046	3.66	13
Parietal Inf R	40	-	52	-42	53	4	.0045	3.55	14
Frontal Sup R	9	-	14	43	41	2	.0045	3.52	15
Angular R	39	-	39	-55	28	3	.0044	3.45	16
Temporal Mid L	38	-	-49	8	-26	2	.0043	3.26	17
Angular R	39	-	33	-68	50	2	.0042	3.22	18
SupraMarginal R	39	-	62	-49	28	2	.0042	3.20	19

Voxel-wise Bonferroni corrected $p=0.05$ corresponds to uncorrected $-\log_{10} p=6.04$.

TABLE 4.7: The SIM-ASN contrast is computed by subtracting group-average voxel-wise r^2 .

The significance is reported by two-tailed Wilcoxon signed-rank test before multi-comparison correction. The cluster is reported only if the average r^2 of SIM is higher than ASN. No cluster-size limit was used when computing connected clusters. Significant small clusters are found in left superior frontal cortex and anterior cingular cortex. Additional near-significant clusters are located in left superior temporal gyrus. No ventroanteriorotemporal cluster is found for SIM.

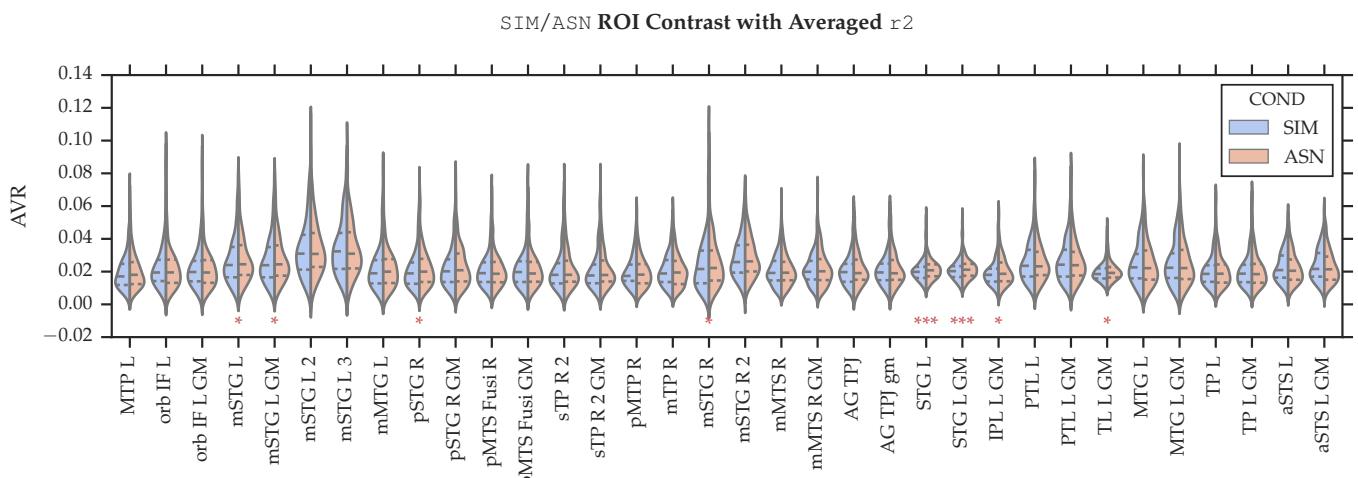


FIGURE 4.11: *: $p<0.05$ uncorrected, **: 0.05 ROI-wise multi-comparison corrected. Red color for ASN.

The average r^2 of voxels in a ROI is computed. We select only ROIs with scores >0.02 in either of SIM and ASN models. ROIs are of minimum size of 26 voxels (radius of 7 mm). None of the tested ROI reveals a significant mean difference in preference for SIM. ROIs in left middle-posterior STG, left inferior parietal lobe and left temporal lobe respond better to ASN model.

SIM-ASN Voxel Contrast, Preference for ASN

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2	$-\log_{10}$	Cluster ID
Cuneus R	18	VisualAssoc	5	-77	22	3	.0083	9.99	1
Cuneus L	18	VisualAssoc	2	-87	25		.0083	8.27	1a
Calcarine L	18	VisualAssoc	2	-83	15		.0070	6.30	1b
Temporal Inf R	20	-	46	-17	-26	2	.0090	9.21	2
Cerebellum 6 L	18	VisualAssoc	-8	-80	-13	4	.0087	8.78	3
Parietal Sup L	7	-	-27	-71	56	2	.0086	8.70	4
Occipital Mid R	18	VisualAssoc	36	-80	3	2	.0078	7.70	5
Cerebellum 8 L	37	Fusiform	-21	-55	-45	3	.0078	7.67	6
Occipital Mid L	19	-	-30	-77	15	2	.0076	7.29	7
Cerebellum Crus1 L	18	VisualAssoc	-2	-80	-16	2	.0075	7.16	8
Fusiform R	18	VisualAssoc	30	-83	-4	3	.0073	6.85	9
Hippocampus R	50	Thalamus	17	-11	-7	2	.0072	6.69	10
Vermis 10	37	Fusiform	-2	-42	-32	2	.0072	6.65	11
Fusiform R	19	-	33	-71	3	2	.0072	6.64	12
Fusiform L	36	Parahip	-33	-26	-19	2	.0071	6.45	13
Thalamus L	50	Thalamus	-2	-20	3	2	.0070	6.36	14
Calcarine L	18	VisualAssoc	-27	-64	6	3	.0069	6.24	15
Calcarine R	17	PrimVisual	17	-83	6	2	.0069	6.22	16
Calcarine R	17	PrimVisual	11	-80	9	2	.0069	6.20	17
Cerebellum 6 R	19	-	14	-64	-13	2	.0068	6.04	18
Occipital Mid R	18	VisualAssoc	30	-93	15	2	.0068	6.02	19
Parietal Inf L	40	-	-30	-39	37	2	.0068	6.01	20
Calcarine R	17	PrimVisual	5	-64	15	2	.0067	5.94	21
Calcarine L	17	PrimVisual	2	-87	6	4	.0067	5.88	22
Occipital Mid L	19	-	-39	-74	6	2	.0066	5.79	23
Calcarine R	17	PrimVisual	11	-68	15	2	.0066	5.79	24
Cerebellum 8 R	37	Fusiform	30	-64	-54	2	.0066	5.76	25
Cerebellum 9 L	37	Fusiform	-11	-42	-32	2	.0066	5.75	26
Calcarine L	17	PrimVisual	-8	-87	3	2	.0065	5.66	27
Lingual L	19	-	-14	-45	-7	2	.0063	5.45	28
Temporal Mid L	20	-	-39	-4	-26	2	.0060	5.21	29
Calcarine L	18	VisualAssoc	-8	-96	-13	2	.0060	5.12	30

Voxel-wise Bonferroni corrected $p=0.05$ corresponds to uncorrected $-\log_{10} p=6.04$.

TABLE 4.8: The SIM-ASN contrast is computed by subtracting group-average voxel-wise r^2 . The significance is reported by two-tailed Wilcoxon signed-rank test before multi-comparison correction. The cluster is reported only if the average r^2 of ASN is higher than SIM. No cluster-size limit was used when computing connected clusters. ASN's model advantage over SIM is often found in bilateral visual association and primary visual areas. Clusters in ventroposterior aspects of temporal lobe is also found in right fusiform, parahippocamal gyri.

Chapter 5

Discussion

5.1 Back to Hypothesis

By modeling *similarity* processing axis by SIM and *association* by ASN features, our hypothesis, which is based on convergent literature findings, argues for a bilateral aTL loci of semantic hub, more precisely located in the ventrolateral aspects. By indicating the presence of a content word, the regression model improved the voxel-models located in bilateral mid temporal pole, left middle fusiform gyrus without differentiating *similarity* and *association*. The constructed SIM features most improved voxels located in bilateral middle temporal gyrus (TG) and superior parietal gyrus, SIG in posterior inferior TG, right angular gyrus and bilateral inf frontal gyrus pars triangularis. The SIM-ASN contrast revealed preferential models for SIM in left superior frontal cortex and anterior cingulum. Near significant results in TL are found in bilateral pSTG, left aMTG, right mMTG. No remarkable contrasts are found in inferior and ventral part of TL. The SIG-ASN contrast found SIG's advantage in primary auditory cortex, posterior fusiform gyrus. The inferior aspect of SIG's positive contrast is primarily located in posterior region of TL.

ASN improvements and contrasts with *similarity* consistently report voxel clusters implicated in associative computation (occipital cortices, angular gyrus (AG), right mMTG, left pMTG and frontal areas), supporting our hypothesis on *association* modeling and localization.

While our findings for SIM/SIG modeling *similarity* are not convergent with the hypothesis, a much simpler construction of CWRATE captured voxel model improvements in ventral aspects of posterior TL and MTP in anterior TL, which are close to the hypothetical loci. This finding leads to multiple arguments to explain such contradictions.

5.2 Precise and Informative Semantic Feature Design

Syntagmatics and Modality-Specificity of Association

Association is proposed as an umbrella term for all non-*similarity* information. Thus ASN embeddings are built as the residual of subtraction of *similarity* embeddings from a mixed embedding. However, since we used GloVe and DepGloVe as our mixed embedding, the corpora used to build these two embeddings are purely textual, thus no explicit perceptual data are provided. An embedding space, composed majorly by syntagmatic information, found its better encoding voxel-model in multiple primary visual and visual association areas (bilateral BA17, 18, 37) and

hippocampal and parahippocampal areas (which are associated with social interactions, episodic memory) when contrasted with *similarity* (No visual area is reported by contrasting ASN with non-embedding features). This finding is convergent with our hypothesis on *association* constructions: modality-dependent, association with episodic memories. Thus the modality-independent aspect of our *similarity* embedding models, which is presumed to extract the rest of information, can also be partially confirmed.

Impact of CWRATE

CWRATE indicates the necessity of semantic retrieval and processing, thus wraps both *similarity* and *association* aspects. Most of the voxel-clusters improved by CWRATE are associated with visual recognition / identification (posterior fusiform), visual association (BA18/19, V2,3,4,5), premotor (Rolandic Oper BA6) visuo-motor coordination (Precuneus, Superior Occipital). But two aTL regions in MTP are also reported, which are near to the neural fiber convergence zone.

As CWRATE is a shared feature for *similarity* and *association* processing, orthonormalizing embedding feature regressors against CWRATE suppresses a large proportion of semantic-axis-specific signal in fMRI encoding, potentially weakening the contrast between SIM and ASN.

Better Constructions of SIM

SIM models have lower regression scores than ASN, this could be due to the semantic-axis processing attribution of voxels, or the lack of quality control of the SIM embedding. The English SIM embedding is well constructed: WordNet is widely used, the resulting embedding's quality is assured by semantic evaluation tasks. Whereas for French the ontology is built upon the algorithm-generated WOLF, which makes use of multilingual resources and is composed of translation-based synsets. The French semantic evaluation task datasets are not tested.

Bullinaria and Levy (2012) found that removing the initial PCs of singular-value-decomposed (SVD) semantic matrices improves the performance on multiple semantic tasks (such as TOEFL, Distance Comparison, Semantic Categorization and Clustering Purity, fMRI encoding/decoding tasks are not included). In our project we did not remove the initial PCs nor did Bullinaria and Levy provide a practical suggestion on the number of PCs to be pruned. The influence of first PCs in obtained SIM is very pronounced, they one-hot encode POS information, so that words are organized by grammatical categories in different linearly dissociable sub-spaces.

There are two aspects of SIM embedding ameliorations. Firstly, unclear on whether the human brain recruits different neural structures for words of different grammatical categories, removing the first PCs of SIM might better approximate the argued *similarity* axis. Secondly, the PC removal is beneficial for computational advantage of model regressions. For example, SIG, is not an embedding resulting from a PCA, thus have no dominant dimensions in the embedding. By promoting voxel-model scores, SIG revealed more voxel clusters in contrast with ASN.

Corpus and Embedding Compatibility

For out-of-vocabulary words in the *Le Petit Prince* (around 5% of the vocabulary), null vectors are used to substitute (unknown) semantic values in this project. The semantic vectors however could be approximated using synonyms or associates available in embeddings. The selection of alternative words should be compatible to the semantic axis of the embedding in question.

5.3 Limits of fMRI

Ventral BOLD Signal Recording

The adopted multi-echo fMRI sequence is adopted to better extract BOLD signals in ventral cortical areas. Traditional fMRI imaging suffers a low signal-to-noise ratio in the region due to the sinuses located near temporal poles, unable to reveal neural activations (J. T. Devlin et al., 2000). The effect of *similarity* and *association* contrast might be subtle. Despite an improved fMRI targeting higher SNR, it could be suspected that the minute contrast could not be shown by fMRI.

Temporal Dynamics of Two Semantic Axes

Ralph et al. (2017) states that in ventroanterior temporal lobe, domain-level semantic distinctions are available around 120 ms post stimulus onset, and around 250 ms detailed semantic information is activated. Kutas and Hillyard (1984) suggest that N400 signal of event-related potentials (ERP) is related to word expectancy and semantic association. Frank and Willems (2017) contrasted word predictability (syntagmatics) and semantic similarity (paradigmatics) with an EEG and fMRI experiment. While the N400 signal is the same for two type of tasks, *similarity* is correlated with mid temporal pole and angular gyrus (consistent with Table 4.6), while word surprisal is correlated with fusiform gyrus (FG), and middle-posterior STG. ASN is aligned with word predictability, however in our experiment it is correlated with FG improvements but not with STG.

The comparison suggests that our fMRI encoding found N400-related signals but not 120 ms and 250 ms neural correlates. While unclear if the earlier activations are masked by later ones, it seems an analysis on temporal axis might be crucial for *similarity* and *association* dissociation. Other imaging techniques other than fMRI (e.g. MEG) might reveal more evidence.

Beyond Lexical Semantics

In a post-hoc analysis, we correlated different subject-wise model scores trained in 9 cross-validations with each participant's multi-choice comprehension questions (Table B.11). Surprisingly, voxel-models' maximum and mean scores of the semantic embedding regressors are negatively correlated with question correct rate. SIM and SIG's correlations are the most extreme ($p < 0.05$ for both max and mean), ASN's mean is near-significance. MIX is not significant but a negative trend is shown.

The findings suggest that [TODO, ANOVA] the semantic understanding modeling should not constrained by lexical semantics. To acquire comprehension of a full phrase, sentence or passage, the brain should transform and integrate each lexicon unit's semantic value. A high semantic model performance might be an indicator of

dominance of mental representations for lexical semantics, and the phrasal integration of semantic values is to a limited extend. The inequivalent contribution of *similarity* and *association* in phrasal semantic processing is also suggested, as *similarity* principle evolves on *absentia* of a lexical unit's occurrence, with *association* pertains a more global view on the present passage.

Again, if lexical semantic information is present post stimulus, it could be better investigated in a tight time window , and the contrast between *similarity* and *association* lexical semantic values might be more clear.

5.4 Statistics

The threshold in whole-brain voxel model performance visualization (r^2 maps) is fixed at 0.005. This choice was arbitrary, and its utility is to filter out uninformative voxels without considering statistical significance of the regression results. As different voxels had different preferential feature dimensions across cross-validation sessions, individuals and models, the group level of significance test for r^2 was a complex question. Since significance for r^2 is not computed for model regression results, the cluster analysis was performed with selections of a certain proportion of best modeled voxels.

Additionally, the F-test result presentation on nested-model improvement contrast is also controversial as it manipulates p-values. The original design was to present a superposition of significance maps of each individual so that regularities could be tracked, and the geometric mean of p-values is equivalent to algorithmic mean of log p-values. The statistical robustness however is questionable.

In this project the contrast of model performances lacks in statistical significance: the voxel-model model contrasts had p-values < 0.001 uncorrected, but none survived voxel-wise multi-comparison correction. A small effect size was foreseen since *similarity* and *association* contrasts are minute. However given the time constraint of the project, recruiting more subjects for additional fMRI recording was not a viable option.

5.5 Cognitive Accounts on Coherence between Embeddings, Semantic Principles and Semantic Hub

We name *similarity* the internal organization of the presumed semantic hub, and argued that *similarity* is principally constituted with paradigmatic axis proposed by Jakobson and Halle. Additionally with pathological evidences, multiple properties of *similarity* axis are defined: cross-modality and conceptual hierarchy. *Similarity* is modeled by embedding based on WordNet-alike ontologies, conformably constructed with the hypothetical properties of the semantic hub. Semantic evaluation tasks based on word-pair proximity evaluation suggest the validity of SIM/SIG model against *similarity* , especially for English embeddings where WordNet and evaluation benchmarks are well founded. Yet no prior experiences confirm the bridging of *similarity* and the semantic hub.

Pathway to Semantic Hub: Accumulative or Differential?

In our hypothesis, we presumed that the semantic hub holds a global view of all representational or operational semantic information. Correspondingly, a holistic *similarity* embedding is constructed, containing all semantic entities. Yet, such construction underestimates the participation of non-hub structures (which are domain-specific or feature-specific). The semantic hub is argued to be graded anatomically based on cytoarchitecture findings (Ralph et al., 2017), an alternative of accumulative information integration from semantic spokes along the processing pathway is a differential semantic information generalization: specific semantic representations are available in non-hub structures, and semantic hub does not represent a copy of information, but computes a more generalized/abstracted version.

Under this hypothetical vision, the *similarity* principle proposed by this project corresponds not only to the strict semantic hub, but also to neural structures connecting to the hub loci. With voxel-specific feature-dimension selection, the regression results might provide insights by analyzing the spatial distribution of preferential effective feature-dimension of voxel-models, if evidences for an increasing semantic specificity is found along the PCA components.

Chapter 6

Conclusion

In this project, we propose two types of semantic embedding spaces encoding respectively semantic *similarity* and *association*. We selected a set of *similarity* semantic relationships and converted semantic ontologies to similarity embeddings. We proposed the usage of general linear model to dissociate *similarity* information from *association*, which are mixed in classic statistical distributed embeddings. Each constructed space contains pure semantic information of one semantic axis, confirmed by semantic evaluation tasks specifically constructed for each axis and example examinations. The collected evidences suggest that the *similarity* embedding construction and GLM dissociation methods are valid.

With built embeddings, we test the hypothesis of aTL localization of semantic hub, which has a hypothetical internal organization by *similarity*. The embeddings are combined with basic features to encode fMRI BOLD signals. While *similarity* embeddings find mostly mid temporal, sup parietal and angular improvements when contrasted with basic features, and sup/mid frontal, a Cingulum, sup/mid temporal with *association* features, *association* models found occipital, frontal, mid temporal, inf tri frontal contrasted with basic features, and inf Temporal, occipital and parahippocampal improvements contrasted with *similarity*. While the results are expected for *association* since associative cortical areas are correlated, the argued aTL loci is not supported by the data.

We argue that the contrast between *similarity* and *association* could be improved with a better construction of *similarity* embeddings and encoding design matrices. Other imaging methods other than fMRI might better show a temporal aspect of the two-axis contrast. Finally, we reviewed our hypothetical structure of *similarity* of the semantic hub, proposing alternative models utilizing contextual local semantic comparison.

Appendix A

Supplementary Methods

A.1 fMRI Stimuli Preparation

A.1.1 Natural Story Stimuli

The following section is translated from Todorovic and Pallier (2018), section 5.2.1.

For the comfort of the participants and their concentration on listening comprehension, the audiobook is divided into 9 blocks, so that each block lasts at most 15 minutes. At the beginning of the French narration of *The little prince : a French/English bilingual book*, the audiobook-related information are not included in the stimuli. For each chapter, the reading of chapter title is removed from the audio, and 3 seconds of silence is added. In Table A.1 the 9-block division is detailed.

A.1.2 Behavior Control

The following section is adapted from Todorovic and Pallier (2018), section 5.2.2, 5.2.3.

As a behavioral control, 4 multiple-choice listening comprehension questions are posed after each story block. The questions are selected and adapted from the English questions used in the fMRI acquisition experiment developed by Cornell University (“Neural Computational Models of Natural Language”, PI: John Hale and Christophe Pallier). Each question is provided with 4 choices.

French <i>The little prince : a French/English bilingual book</i> Chapter Division			
Block	Chapters	Duration	fMRI Images
1	1-3	10:12	309
2	4-6	10:48	326
3	7-9	11:43	354
4	10-12	10:25	315
5	13-14	09:41	293
6	15-19	12:31	378
7	20-22	10:59	332
8	23-25	09:44	294
9	26-27	11:08	336

TABLE A.1: fMRI TR=2s.

To control for the difficulty of the questions and to ensure that participants must have firstly attentively listened to the story to successfully respond to the question, the same questions are distributed and tested to French native speakers without a priori exposure to *Le Petit Prince* in the last 5 years via Information Relay in Cognitive Sciences¹. The collected responses are used as a control group to test if the fMRI participants respond significantly better.

Additional open comprehension questions are asked to engage the participants into short conversations during the fMRI recording. The questions are asked orally, with a visual aide of a sampled drawing from the currently-played block chapters. If the participants corrected answers three comprehension questions, they are asked to retell the passage concerning the presented image.

All questions are available in the annex of Todorovic and Pallier (2018).

A.2 fMRI Acquisition

The following section is translated from Todorovic and Pallier (2018), section 5.1.

Subjects

Continuing from Section 3.1, the recruited subjects have not exposed to the story of *Le Petit Prince* for at least 5 years, including books, audiobooks and films. They should not have a clear memory of the story.

Experiment Procedure

A Siemens MRI scanner at 3 Tesla acquires fMRI images when the participants passively listen to a narration. Each fMRI recording session lasts at most 90 minutes for security considerations, so the 9 blocks of story is recorded in two sessions within the same day, with a 60 - 90 minutes break between. One session consists of 4 or 5 blocks.

The participants were invited half an hour before the start of the MRI acquisition to have an interview with the Neurospin medical doctor. After the interview, they were received by Todorovic and Pallier. The receiver orally explained the procedure. Then the participants were placed in the scanner for an anatomical acquisition session. This session lasted 8 minutes, during which the instructions were displayed, from a pdf file (available in annex of Todorovic and Pallier (2018)), on a screen that could be seen through the mirror that was attached to the participants' head. After the instructions, the images appearing in the first two chapters of the Little Prince were presented, since they were relevant images for understanding the story. After the anatomical acquisition, a sound test (despite the MRI acquisition noise) was performed by playing the introductory sentence of the audiobook. This audio was chosen for the sound test because it had similar acoustic properties to the rest of the audio book and was not used during the listening afterwards. When the sound level was adjusted, text listening starts. Participants listened to the text with their eyes closed to prevent eye movements from disrupting the BOLD signal.

¹Relais d'information en sciences de la cognition, <https://www.risc.cnrs.fr/>.

After each block of text, the participants opened their eyes and answered the comprehension questions, displayed on the screen one by one. After reading the question and the proposed answers, the participant gave his answer orally via the intercom and the experimenter recorded the answer. When the given response was not easy to distinguish ("b" or "d"), the participant was asked to read the beginning of the chosen response or to give a word that begins with the letter associated with the chosen response (for example, "b" as a banana"). After the comprehension questions, the open questions were asked, or the participant was asked to retell the heard story. The answers to the open comprehension questions were recorded by the microphone on a mobile phone. A myopia participant did not answer the reflexion questions since she could not see the images without glasses, which she could not wear inside the scanner.

At the end of the story, an additional five-minute fMRI acquisition was performed. The participant listened to sentences in French and unintelligible audio stimuli, obtained by acoustic deformation of the sentences in French. This procedure allows the language processing areas in the participant's brain to be quickly located.

A.3 Regression Parameters

The tested α values are 0, 10^n for n in 0, 0.6, 1.2, 1.8, 2.4, 2.5, 2.58, 2.66, 2.74, 2.82, 2.89, 2.97, 3.05, 3.13, 3.21, 3.29, 3.37, 3.44, 3.53, 3.61, 3.68, 3.76, 3.84, 3.92, 4, 4.01, 4.5, 5, 5.50, 5.99, 6, 7, 8.

The tested feature dimensionalities are 1 (RMS), 2 (WRATE), 3 (CWRATE), 4 (begin of embedding features), 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 25, 35, 45, 55, 65, 75, 85, 95, 103 (end of SIM), 105, 115, 125, 135, 145, 155, 165, 175, 185, 195, 203.

A.4 Supplementary Analysis

A.4.1 Non-nested Model Comparison

Following the original pipeline proposed in Merkle et al., 2016, non-nested model comparison should first test for non-equivalence, then for distinguishability, then compare model performance.

The particular case for SIM and ASN model comparison partially validate the non-equivalence, since the regressor bases are constructed by linear de-correlation, of which the objective is to maximize the found co-linearity between SIM and MIX spaces, thus minimize that between SIM and ASN.

Despite the non-equivalence of semantic models, co-linearity could be introduced in the regressor building stage, where a convolution is introduced. For the distinguishability, we proceed by using the constructed regressors. We try to test the distinguishability in this particular sample of two semantic representation spaces (against the fMRI stimuli's text data, with application of a convolution filter), by performing linear regressions between the two design matrices (as a collection of regressors).

To simplify the conceptual construction, we proceed similarly with fMRI encoding: from the 9 design matrices of one semantic space, we iteratively leave out one as validation data, the other 8 being training data. We use training data to learn a GLM

mapping between different semantic embeddings. Then we test the generalization performance of the predicted model on the validation data.

The comparison is two-fold: the first using SIM to predict ASN, the second in the opposite direction.

A.4.2 Comprehensive ROI List

The collected ROI peaks from literatures are available at http://bit.ly/micipsa_roi_list.

Appendix B

Supplementary Results

B.1 Semantic Embeddings

B.1.1 Visualization of Semantic Spaces with TensorFlow Projector

The French SIM space's first four PCs encodes POS information. Thus the 3D PCA presentation of SIM is three axes parallel to the visualization axes (Figure B.1).

B.1.2 Semantic Ranking Task Results

B.1.3 Vocabulary Coverage by POS

B.1.4 Corpus-Targeted Semantic Feature Selection

[TODO still necessary?]

B.2 Non-nested Model Comparison

Use SIM to Predict ASN

[TODO, insert heat-map of coefficients] Each of the 203 columns in the ASN class design matrix (including non-semantic features) are predicted by 103 columns of the SIM class design matrix. We averaged the r^2 score of each column model across 9 cross-validation runs. The histogram of the scores are plotted in Figure B.4, informative model scores are presented in Table B.4. As our design-matrices are orthonormalized, columns sitting at larger indexes have a dependency on smaller-indexed columns. The first columns being bell predicted start at index 12 (to 14). Other columns are scattered up to index 47. We can therefore conclude that the predictability found are purely due to numerical coincidences.

Use ASN to Predict SIM

The same procedure yields also 14 effective column models for SIM. The correlation coefficients are significantly higher than the models predicted with SIM matrices. Besides, the first 5 columns of SIM are all well predicted ($r > 0.30$) by ASN, indicating there's partial signal information overlap between the two models. Since Section B.1.1 shows that the first 4 dimensions in SIM one-hot encode POS information, it is reasonable that POS information is also traceable in syntagmatic-information dominated semantic embedding.

To further investigate the column-wise correlation, we also plot the coefficients of each ASN column learned by GLM. [TODO heatmap average].

French Semantic Embeddings

Target word: professeur_n

SIM	SIG	ASN	MIX
pédagogue_n	pédagogue_n	fondateur_n	naissance_n
éducateur_n	éducateur_n	psychose_n	psychose_n
instituteur_n	instructeur_n	éducation_n	éducation_n
instructeur_n	instituteur_n	serveur_n	secrétaire_n
arbitre_n	arbitre_n	secrétaire_n	logique_a
lecteur_n	adjudant_n	défenseur_n	fondateur_n
enseignant_n	passe_n	imitation_n	chronique_a
passee-partout_n	passee-partout_n	vicaire_n	imitation_n
passepourtout_n	lecteur_n	sensation_n	sensation_n
passe_n	enseignant_n	protecteur_n	honneur_n
adjudant_n	abonné_n	protectionnisme_n	traumatisme_n
aide_de_camp_n	maestro_n	volontaire_n	vicaire_n
maître_n	spécialiste_n	fonctionnaire_n	facilité_n
maestro_n	châtelain_n	naissance_n	serveur_n
capitaine_de			
_vaisseau_n	capitaine_n	photographie_n	proposition_n
maître_d'hôtel_n	commandant_n	producteur_n	disparition_n
commandant_n	propriétaire_n	évidence_n	moteur_n
capitaine_n	professionnel_n	honneur_n	sagesse_n
spécialiste_n	leader_n	croisade_n	croisade_n
commandement_n	contributeur_n	missionnaire_n	évidence_n
abonné_n	possesseur_n	moteur_n	quantité_n
overlord_n	savant_n	pluralisme_n	défaillance_n
châtelain_n	commandement_n	sagesse_n	édition_n
précepte_n	acquéreur_n	coexistence_n	défenseur_n
fondateur_n	acheteur_n	objectif_a	volontaire_n
débutant_n	participant_n	facilité_n	protecteur_n
tyro_n	officiel_n	disparition_n	pluralisme_n

TABLE B.1

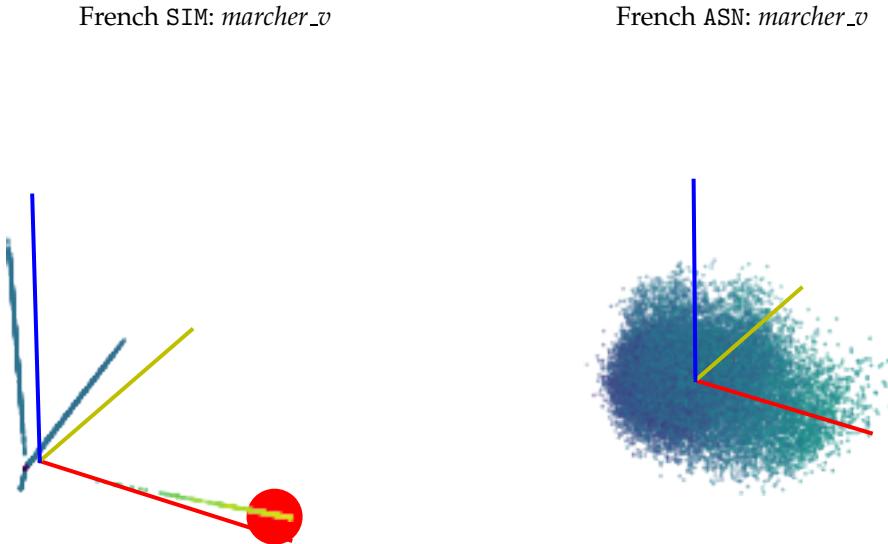


FIGURE B.1: Due to the nature of WOLF, SIM's first PCs denotes POS category. Light colors indicate the proximity of the represented words with *marcher_v*.

FIGURE B.2: ASN's variance are more homogeneously distributed over PCs.

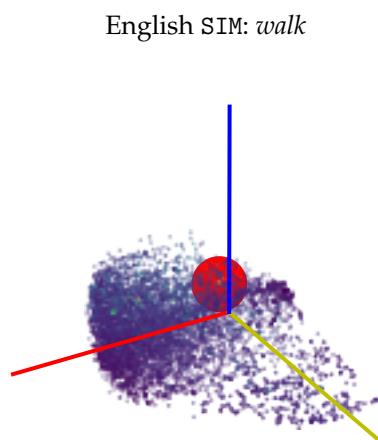


FIGURE B.3: English's SIM however seems not to be macroscopically influenced by POS information.

English WordNetEmbedding Iterations				Association			
Relations	Lexicon	Dim.	Metric	Similarity			
				SIMLEX-999	WS353-SIM	RG1965	WS353-ASN
Synonymy	15k	Pearson	.2256	.2679	.3627	.123	
		Spearmann	.2001	.2003	.3403	.0971	
	60k	Pearson	.234	.2112	.3394	.1449	
		Spearmann	.1747	.1895	.2629	.1129	
S+Antonymy	850	Pearson	.1534	.2743	.373	.0969	
		Spearmann	.1255	.1922	.3302	.0817	
S+Hyper/Hyponymy	15k	Pearson	.3904	.4825	.6187	.0373	
		Spearmann	.4018	.3856	.5145	.0259	
SHeHo+ adj.participle_of_verb+ adj.similar+ adv.derive_from_adj	60k	Pearson	.5079	.5333	.6784	.0525	
		Spearmann	.4986	.4214	.576	.0272	
	15k ⁰	Pearson	.5268	.5483	.6991	.1092	
		Spearmann	.5152	.4757	.5501	.0515	
All ¹	15k ⁰ 511	Pearson	.506			.0279	
		Spearmann	.4989			.0193	
	13k ²	Out Of Vocabulary		.002	.02	0	.012
		Pearson	.5	.65	.65	.32	
Synonym Database ⁴	60k ³	Spearmann	.52	.67	.75	.33	
		Pearson	.5	.51	.56	.31	
	850	Spearmann	.51	.58	.72	.3	
		Pearson	.6814	.5819	.8155	.317	
Database ⁴	60k ⁵	Spearmann	.6566	.4677	.7032	.3153	
		Out Of Vocabulary		.066	.227	.077	.19

⁰ Version reported in Table 4.1.¹ Data reported by Saedi, Branco, António Rodrigues, and Silva (2018).² Cue words selected from psycholinguistic experiment datasets.³ Words selected randomly. On the contrary, our implementation selects the top 60k most frequent words in WordNet.⁴ Synonym database is created by thesauri fusion and symmetrization. Data provided by Ploux and Ji (2003).⁵ The synonym database contains multi-word phrases, whereas task benchmarks only test for single-word pairs. The actual lexicon size of the database is 36718.

TABLE B.2: Placeholder

The Little Prince Vocabulary Coverage

# Instances in fMRI Recording Session									
		R1	R2	R3	R4	R5	R6	R7	R9
Nouns									
Story	T	269	286	306	242	284	355	281	265
	V	142	140	152	107	108	147	109	121
56665	TM	17	3	7	5	14	14	11	11
	%	6.32	1.05	2.29	2.07	4.93	3.94	3.91	4.15
	VM	10	3	6	3	3	8	6	7
	%	7.04	2.14	3.95	2.80	2.78	5.44	5.50	5.19
ASN	TM	20	9	8	10	22	16	11	13
/MIX	%	7.43	3.15	2.61	4.13	7.75	4.51	3.91	4.91
/SIG	VM	12	5	7	4	7	10	6	8
24519	%	8.45	3.57	4.61	3.74	6.48	6.80	5.50	6.61
Verbs									
Story	T	227	274	313	306	258	296	331	278
	V	104	109	142	119	84	113	99	110
56665	TM	9	15	14	15	10	4	7	12
	%	3.96	5.47	4.47	4.90	3.88	1.35	2.11	4.32
	VM	7	7	9	11	9	4	5	8
	%	6.73	6.42	6.34	9.24	10.71	3.54	5.05	7.27
ASN	TM	9	19	15	17	10	8	9	12
/MIX	%	3.96	6.93	4.79	5.56	3.88	2.70	2.72	4.32
/SIG	VM	7	10	10	13	9	7	6	8
24519	%	6.73	9.17	7.04	10.92	10.71	6.19	6.06	7.27
Adjectives & Adverbs									
Story	T	229	252	241	214	190	251	207	169
	V	102	111	117	103	100	107	94	104
56665	TM	10	12	11	7	6	15	6	7
	%	4.37	4.76	4.56	3.27	3.16	5.98	2.90	4.14
	VM	9	6	7	6	4	7	5	4
	%	8.82	5.41	5.98	5.83	4.00	6.54	5.32	4.82
ASN	TM	19	19	15	10	16	36	15	12
/MIX	%	8.30	7.54	6.22	4.67	8.42	14.34	7.25	7.10
/SIG	VM	11	11	9	8	10	15	8	6
24519	%	10.78	9.91	7.69	7.77	10.00	14.02	8.51	7.23
									6.73

TABLE B.3: T: Token, V: Lexicon Unit, M: Miss

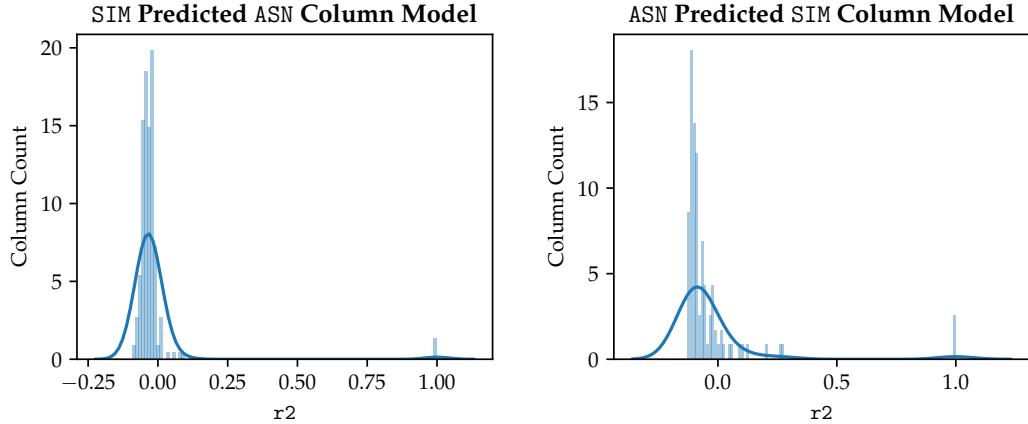


FIGURE B.4: 3 columns are perfectly predicted as they are non-semantic features that are shared by the design matrices. For the rest most of the column-models are worse than arbitrary models.

FIGURE B.5: Most of the column-models are worse than arbitrary models. More columns in SIM are better predicted with ASN design matrices.

SIM Predicted ASN Column Model Performances

Col Index	22	20	14	37	47	18	27	26	13	12	30
r2	.0891	.0834	.0548	.0346	.0164	.0152	.0122	.0113	.0112	.0075	.0004
r	.2984	.2887	.2340	.1861	.1279	.1233	.1104	.1064	.1058	.0867	.0203

TABLE B.4: Index starts at 0, ASN group features starts from 3. Among the 14 informative models ($r^2 > 0$), 3 are non-semantic features (not listed above). Pearson's r are converted from r^2 .

ASN Predicted SIM Column Model Performances

Col Index	6	3	5	11	4	7	75	18	42	25	47
r2	.2761	.2646	.2041	.1256	.1032	.0971	.0529	.0485	.0192	.0158	.0107
r	.5254	.5144	.4518	.3544	.3213	.3116	.2299	.2202	.1385	.1256	.1037

TABLE B.5: Index starts at 0, SIM group features starts from 3. SIM columns are much better predicted by ASN.

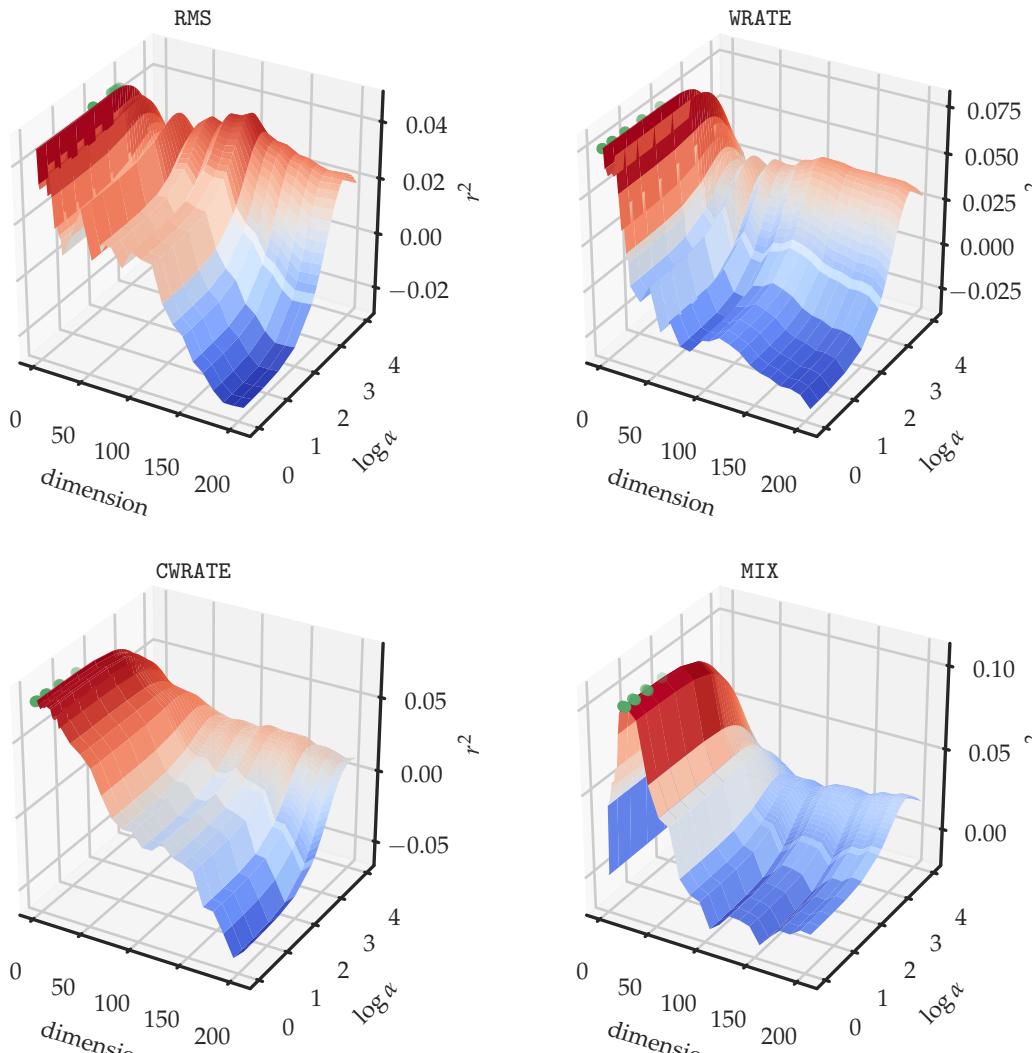


FIGURE B.6: A selection of typical voxels found in MIX model regression results of subject 1 run 1. As shown by all four voxels, the regularization by large α is beneficial only in higher feature dimensions. Our selection of α contains a near-optimal value candidates for these voxels since the curve is all declining for the largest values. Green plots indicate the top 10 best configurations of the given voxel.

B.3 Regression

B.3.1 More on α and Dimension Selection

Completeness of Research Space

For illustrative purpose, we selected four typical voxels in post-hoc from the regression results of MIX, run 1 subject 1 (Figure B.6)¹. MIX is used since it is the default semantic space used in other works and has not been modified. Each voxel is the best modeled voxel who maximizes r^2 among all α s using only partial feature information of a certain regressor class. For example, a CWRATE class typical voxel is a voxel of which the best r^2 is achieved with *all* first 3 features ($\text{dim}=3$).

¹Interactive version of the plot available online <https://plot.ly/~neegola/11/>.

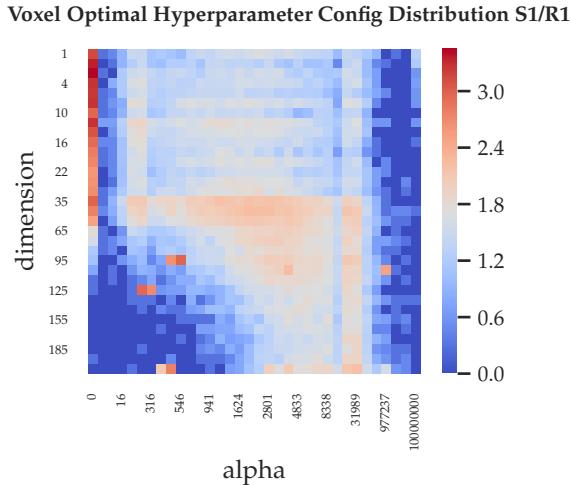


FIGURE B.7: [TODO fix title display] The scores are retrieved from MIX model regression for subject 1, fMRI run 1. Each cell represents an α and dimension combination. The color indicates the logarithm of number of voxels having its global optimality with a given parameter set after having filtered out non-informative voxels ($r^2 < 0$). For small dimensions (< 35), small α s (including 0) achieve the best performance. Starting from dimension 35, Ridge regularization with larger α s is necessary.

The optimal r^2 s of typical voxels are never attained at the upper bound of α and dimension space. We plotted the heat-map for all voxels from the same run to verify if it is also the case at the whole-brain level (e.g. Figure B.7). We averaged 9 cross-validation run results to visualize subject-level best configuration distribution (Figure 4.4). Histograms of best dimension and α voxel-configuration of the averaged results are also plotted in Figure B.8. Plots for all runs and all subjects are available online². The analysis confirms that our parameter combination test range contains the near-optimal configuration for each voxel.

α Variability across Voxels

[TODO: histogram animation with alpha evolution,]

[TODO: discussion on overfit by dimension, despite regularization]

B.3.2 Impact of CWRATE

[TODO: Taking over SIM/ASN ? Visualize regression results without CWRATE]

B.4 Embedding Model Brain Maps

B.4.1 SIM and ASN F-test

B.4.2 SIG Nested Model and Contrast

[TODO: replace SIM by SIG]

[SIG 209, No particularly different results are found. in PeaksSIG.ipynb]

²[TODO] add public url

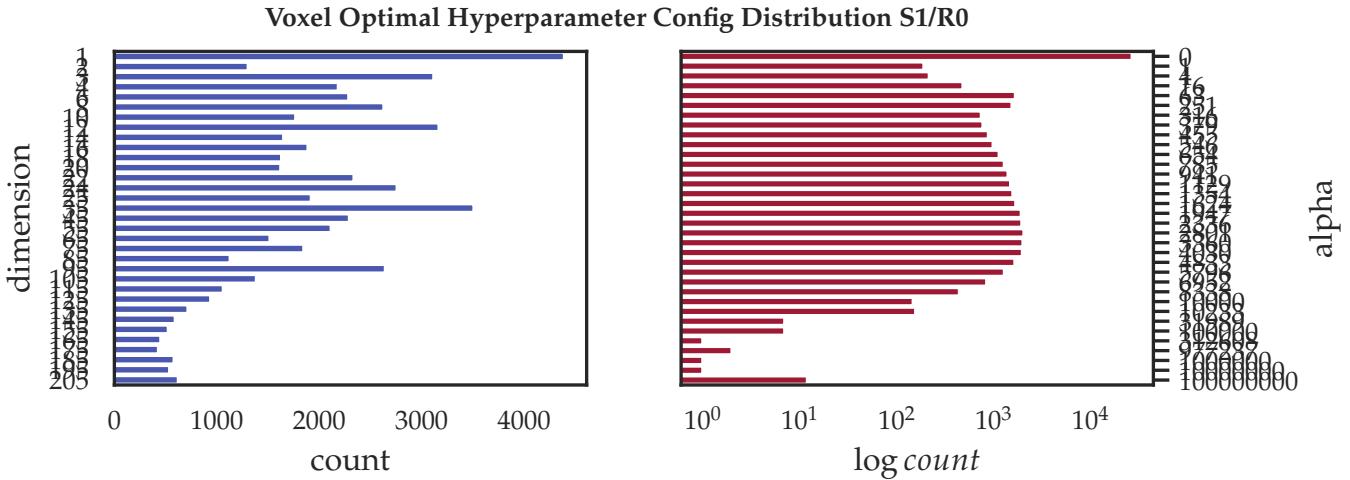


FIGURE B.8: [TODO, fix axis display] **Left panel:** Among informative voxels, a large portion of voxels are best modeled by RMS and CWRATE regressor classes. The addition of each semantic dimension from MIX improved a small proportion of voxel-models, marginal might be the contribution. **Right panel:** Most of the voxels are best modeled without Ridge regularization. The larger number obtained at $\alpha = 10000$ might indicate that larger alpha values might help better model a thousand voxels. A cumulation of voxel-count at the upper bound of the α axis is noted: we performed post-hoc test for larger α values than in the initial configuration, but the improvement of r^2 over the original score is marginal ($< 10^{-4}$) for a sample of these voxels. A post-hoc analysis of larger α s indicates a limited improvement of r^2 , thus for computational simplicity we kept the original Grid Search space of α .

ASN Best Improved Voxel Clusters

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2 Peak	$-\log_{10} p\text{-value}$
Top 4.7%							> .0065	>4.18
Temporal Mid L	37	Fusiform	-53	-58	-3	131	.0121	4.35
Frontal Inf Tri L	46	-	-46	37	11	68	.0124	4.35
Occipital Mid L	39	-	-32	-78	41	242	.0125	4.35
Frontal Sup Medial L	9	-	-5	58	32	32	.0092	4.35
Cingulum Mid L	31	-	-1	-36	43	34	.0105	4.35
Frontal Mid R	8	-	28	15	46	63	.0093	4.35
Angular R	39	-	43	-75	39	191	.0114	4.35
Temporal Mid R	21	-	46	-32	0	182	.0106	4.35
Frontal Inf Tri R	46	-	50	35	8	56	.0107	4.35

TABLE B.6: The Wilcoxon signed-rank test's p-value is thresholded at $10^{-4.18}$ to make a clean cut is found in p-value histogram. Only 13 voxels reach a significance level of $10^{-4.35}$ as is done with ASN. The voxel selection leads to top 4.7% important voxel-model improvements. The largest r^2 boost is found in left MOccipital, followed by right mSTS, right Angular, left pFusiform bilateral iTriFrontal. Relatively smaller clusters are found in MCingulum L, L S Frontal BA9, R M Frontal BA8.

CWRATE/SIM/SIG/ASN F-test Significant Voxels

Position	BA	Functional Label	x	y	z	# Voxel in Region
CWRATE						
Fusiform L	19	-	-30.0	-71.0	-13.0	1
Occipital Mid R	19	-	46.0	-77.0	3.0	1
Cerebellum 8 R	37	Fusiform	24.0	-42.0	-45.0	1
Cerebellum 9 R	37	Fusiform	8.0	-52.0	-35.0	1
Lingual R	18	VisualAssoc	8.0	-90.0	-7.0	1
Precuneus R	7	-	5.0	-55.0	75.0	1
Cingulum Post R	30	-	2.0	-33.0	12.0	1
Occipital Sup L	7	-	-24.0	-83.0	44.0	1
Fusiform L	18	VisualAssoc	-24.0	-83.0	-16.0	1
Lingual L	18	VisualAssoc	-24.0	-90.0	-16.0	1
Cerebellum Crus2 R	37	Fusiform	46.0	-74.0	-42.0	1
Occipital Inf R	37	Fusiform	52.0	-68.0	-13.0	1
Occipital Inf L	18	VisualAssoc	-30.0	-87.0	-4.0	1
Cerebellum Crus1 L	18	VisualAssoc	-33.0	-87.0	-26.0	1
Cerebellum 8 L	20	-	-43.0	-52.0	-51.0	1
Cerebellum Crus2 L	37	Fusiform	-43.0	-64.0	-38.0	1
Cerebellum Crus2 L	37	Fusiform	-46.0	-61.0	-45.0	2.0
Occipital Mid L	19	-	-49.0	-74.0	-0.0	1
Temporal Pole Mid L	38	-	-52.0	11.0	-35.0	1
Cerebellum Crus2 L	37	Fusiform	-52.0	-52.0	-42.0	1
Temporal Inf L	37	Fusiform	-55.0	-45.0	-23.0	1
Cerebellum Crus1 L	19	-	-30.0	-71.0	-26.0	1
Cerebellum 8 R	37	Fusiform	36.0	-55.0	-51.0	1
Temporal Pole Mid L	38	-	-55.0	11.0	-32.0	1
Rolandic Oper L	6	-	-62.0	2.0	9.0	1
SIM						
Temporal Mid L	39	-	-65.0	-52.0	3.0	1
Heschl L	4	PrimMotor	-65.0	-11.0	9.0	1
Temporal Mid L	21	-	-58.0	-49.0	9.0	1
Temporal Mid L	39	-	-55.0	-55.0	3.0	1
Temporal Sup R	22	-	43.0	-39.0	12.0	1
SIG						
Rolandic Oper R	40	-	55.0	-14.0	12.0	1
Temporal Sup R	22	-	58.0	-30.0	3.0	1
Temporal Sup R	40	-	62.0	-20.0	12.0	1
ASN						
Lingual R	19	-	21.0	-64.0	-0.0	1
Occipital Mid R	18	VisualAssoc	33.0	-80.0	-0.0	1

TABLE B.7: The most severe voxel score selection of RMS leads to left primary cortex (BA41) activation. Also well modeled voxels are distributed in more extensive areas of bilateral BA41 and right BA23 and BA10. With the addition of CWRATE features, voxel performances are systematically improved. With CWRATE, no other clusters appear in the thresholded voxel set. Left BA41 has a higher concentration of best modeled voxels, while right BA41 and cingulum mid R [TODO labels] degrade in voxel score ranking. Right BA10 also improves in ranking.

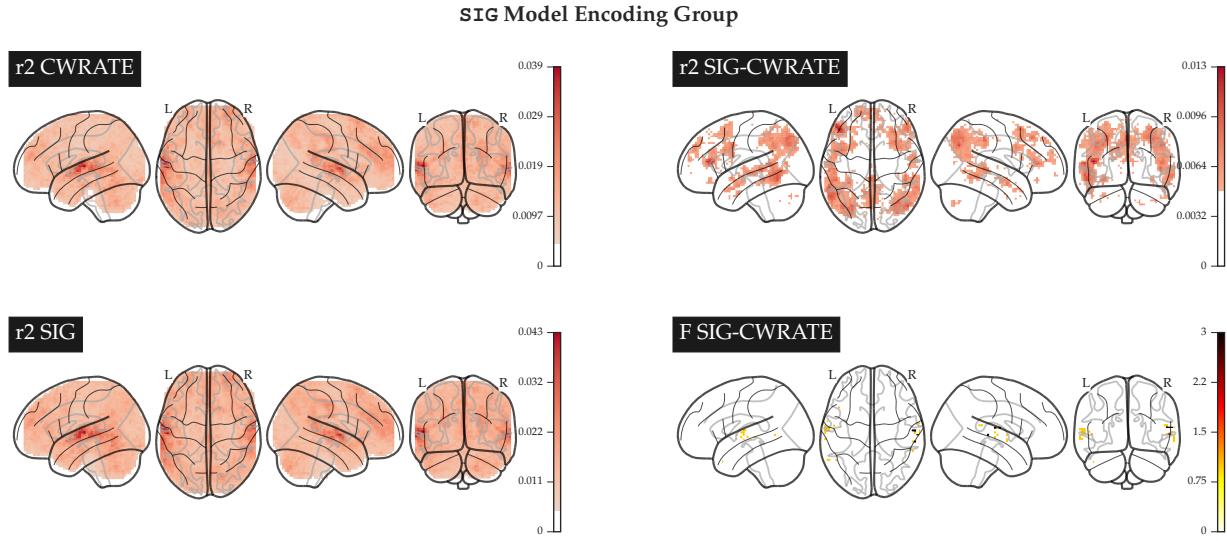


FIGURE B.9: [TODO] **Left panels:** The global activation pattern is unchanged with the feature addition. Best modeled zones are bilateral primary auditory cortices. **Right upper panel** shows that SIM better models bilateral MTC, sup Parietal, Angular Gyrus (part of Wernicke's area), supramarginal gyrus and prefrontal areas (Table 4.6). F-test in **right lower panel** reports significant voxels in left pMTG BA21, 39, right pSTG BA22 and left Heschl BA4 (Table B.7). Subject-wise results are available online at http://bit.ly/micipsa_sim_wholebrain.

SIG Best Improved Voxel Clusters

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2	$-\log_{10} p\text{-value}$
Top .5%								
Temporal Inf L	37	Fusiform	-56	-56	-12	85	.0104	4.35
Frontal Inf Tri L	46	-	-46	35	12	35	.0128	
Parietal Sup L	7	-	-27	-73	44	27	.0093	
Angular R	39	-	49	-70	35	89	.0105	
Frontal Inf Tri R	46	-	48	38	9	17	.0088	

TABLE B.8: [TODO] We thresholded Wilcoxon signed-rank test's significance at $10^{-4.35}$ as a clean cut is found in p-value histogram, which leads to a selection of top .5% important voxel-model improvements. The largest and most improved voxel-cluster is found in left BA21, then in right Angular Gyrus which is part of Wernicke's area (language, V. S. Ramachandran, and Edward Hubbard published a paper in 2003 metaphores). A more lateral and smaller-cluster improvement is found in right MTG.

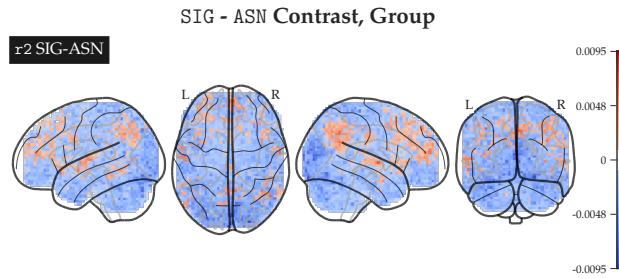


FIGURE B.10: [TODO]The differences of best voxel-model r^2 s are plotted. SIM preference is found in left BA10 SPFC, l aCC, l STS, l MPFC, r IParietal, r STG, MTG (Table 4.7) ... ASN preferences are found in bilateral BA18, RBA2-, l BA7 SParietal, R bA18, l BA37, l BA19, (visual association, fusiform, primary visual, Parahip, and Thalamus). Subject-wise results are available online at http://bit.ly/micipsa_sig_asn_contrast

[TODO, SIM->SIG] We added with upon non semantic-embedding models SIM features to construct *similarity* semantic models. While the whole-brain activation pattern stays globally unchanged (Figure 4.7 for group-wise average), in SIM voxel-models, left primary cortex are better ranked than in BASE model, while right mid cingulum models degrade (Table 4.4). SIM enlarges the performance superiority of left STG over right STG, indicating a left preference for textual semantic *similarity* processing. The shrinkage of Mid Cingulum's proportion in top 1% voxel models might imply that it has a limited participating in *similarity* processing. The r^2 distribution analysis (Figure 4.8 left) shows that in group-average SIM is informative for most of the voxel-models and none of voxels is overfitted by this addition. Table 4.6 reports the most improved voxel clusters by SIM to be located in bilateral MTG, left Sup Parietal and right Angular Cortex ($W=210$, $\Delta r^2 > 0.0079$, $p\text{-value} < 10^{-4.35}$ uncorrected). Left MTG improvements are more extensive and more important than right MTG. F-test results shows that SIM significantly improves isolated voxels (Table B.7, $p\text{-value} < 0.05$ voxel-wise multi-comparison corrected) in left pMTG BA21, 39, right pSTG BA22.

With SIG

Section B.2 suggests that first feature dimensions of SIM can be partially recovered by ASN model. Therefore, ASN might also be able to model voxels using less than 5 features from SIM, the result might thus lack low-level SIM/ASN contrast. As the first 4 dimensions of SIM encodes primarily POS information (Section B.1.1), we performed ad-hoc regressions on SIM space but uses only lemmas from a certain grammatical category to identify possible impacted regions. [TODO, supplementary]

The results found are consistent with the conjectures above: ASN scores are higher than SIM in average (Figure 4.8 right), most of voxels respond better to ASN models (Figure 4.10). As the Wilcoxon test shows ($W=$, $\Delta r^2 > 0.0068$, $p\text{-value} < 0.05$ voxel-wise multi-comparison corrected), only two significant clusters are found for SIM in ... (Table 4.7) and 17 are found for ASN (Table 4.8) in

The reported clusters for SIM are composed of 4 to 5 voxels. In our ROI analysis, ROIs larger than 26 voxels are used, thus none of the ROI revealed significance for SIM. As ASN has an overall dominance for almost all brain regions, small ROIs located in

SIG-ASN Voxel Contrast, Preference for SIG

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2	$-\log_{10}$ p-value	Cluster ID
Angular R	39	-	39	-55	28	8	6.74	.0078	1
Precuneus R	31	-	8	-55	41	13	6.41	.0075	2
Precuneus R	23	-	8	-55	25	4	6.15	.0073	3
Cingulum Ant L	32	-	-2	30	31	14	5.94	.0072	4
Cingulum Ant L	32	-	-11	34	22		4.06	.0054	4a
Frontal Mid L	10	-	-36	49	15	3	5.69	.0070	5
Temporal Sup R	41	PrimAudi	49	-23	6	5	5.45	.0068	6
Precuneus L	23	-	-2	-49	37	14	5.38	.0067	7
Cingulum Post L	23	-	-8	-49	31		3.81	.0052	7a
Caudate R	48	Caudate	17	21	3	4	5.38	.0067	8
Frontal Sup Medial R	10	-	5	56	15	4	5.38	.0067	9
Temporal Sup L	41	PrimAudi	-58	-14	6	8	5.30	.0066	10
Precentral R	8	-	39	8	44	2	5.28	.0065	11
Frontal Sup L	10	-	-27	59	25	2	5.22	.0065	12
Precentral L	6	-	-49	2	28	2	5.03	.0063	13
Temporal Sup L	22	-	-55	-1	-7	2	4.98	.0062	14
Temporal Inf L	37	Fusiform	-49	-42	-13	2	4.77	.0060	15
Temporal Sup R	41	PrimAuditory	62	-8	3	4	4.70	.0060	16
Cingulum Mid R	23	-	5	-20	34	6	4.62	.0059	17
Insula L	47	-	-36	18	-13	2	4.61	.0059	18
Frontal Sup Medial R	8	-	11	34	53	2	4.57	.0059	19
Angular R	39	-	36	-64	41	3	4.57	.0059	20
Caudate R	48	Caudate	14	5	18	2	4.46	.0058	21
Parietal Sup L	7	-	-14	-71	50	4	4.46	.0058	22
Temporal Mid L	19	-	-52	-68	6	3	4.29	.0056	23
Frontal Inf Orb L	47	-	-46	21	-7	2	4.26	.0056	24
Parietal Inf R	40	-	46	-42	41	2	4.26	.0056	25
Precentral R	8	-	39	5	50	3	4.24	.0056	26
Cuneus L	19	-	-14	-74	34	5	4.05	.0054	27
Frontal Sup Medial R	10	-	5	62	12	3	4.01	.0054	28
Cingulum Mid L	23	-	-5	-14	28	3	4.00	.0054	29
Cingulum Ant R	32	-	5	43	3	3	3.94	.0053	30
Precuneus L	7	-	-14	-68	31	2	3.83	.0052	31
Angular L	39	-	-46	-64	47	2	3.81	.0052	32
Frontal Sup Medial R	10	-	2	53	6	3	3.79	.0052	33
Frontal Sup Medial R	10	-	8	59	22	5	3.75	.0051	34
Cingulum Mid L	23	-	-5	-23	34	3	3.74	.0051	35
Precuneus L	7	-	-17	-61	66	2	3.73	.0051	36
Frontal Sup Medial L	9	-	2	56	34	2	3.65	.0050	37
Frontal Sup R	8	-	21	37	53	2	3.40	.0047	38
Parietal Inf R	40	-	43	-45	44	2	3.27	.0046	39

TABLE B.9: [TODO]

SIG-ASN Voxel Contrast, Preference for ASN

Position	BA	Functional Label	x	y	z	# Voxel	Δr^2 Peak	$-\log_{10}$ p-value	Cluster ID
Calcarine L	17	PrimVisual	2	-87	6	26	-	10.26	1
Calcarine R	17	PrimVisual	5	-80	12		.0095	8.30	1a
Cingulum Mid R	32	-		17	5	37	2	.0078	6.81
Calcarine L	18	VisualAssoc	-2	-99	12	2	.0075	6.27	3
Frontal Inf Orb R	47	-		30	27	-23	2	.0065	4

TABLE B.10: [TODO]

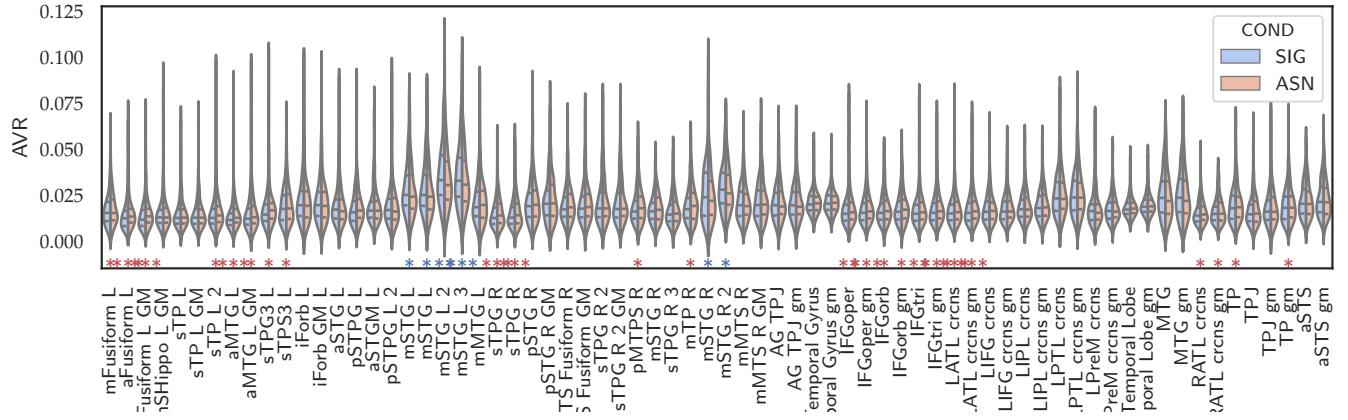
SIG/ASN ROI Contrast with Averaged r^2 

FIGURE B.11: *: 0.05 uncorrected, **: 0.05 ROI-wise multi-comparison corrected. Red color for ASN.

The average r^2 of voxels in a ROI is computed. We select only ROIs with scores >0.02 in either of SIM and ASN models. ROIs are of minimum size of 26 voxels (radius of 7 mm). None of the tested ROI reveals a significant mean difference in preference for SIM. ROIs in left (m,p)STG, 1 IParietalLobe, ITemporal Lobe respond better to ASN model.

left m,p STG and large anatomical structures including IParietalLobe and Temporal Lobe all revealed their preference for ASN model.

B.5 On Behavioral Control

The correlation between cross-validation session-wise model performances consistently correlates with participants' comprehension question scores in the end of each validation fMRI recording.

Correlation of Model Performance with Comprehension Question Scores

	SIM		SIG		ASN		MIX	
	max	mean	max	mean	max	mean	max	mean
Pearson's r	-.15	-.17	-.16	-.18	-.09	-.12	-.11	-.14
p	.0466	.0221	.0262	.0173	.1329	.0593	.4296	.3893

TABLE B.11

Bibliography

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8. doi:10.3389/fninf.2014.00014
- Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Paca, M., & Soroa, A. (2009). A Study on Similarity and Relatedness Using Distributional and WordNet-based Approaches. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 19–27). NAACL '09. Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved November 20, 2018, from <http://dl.acm.org/citation.cfm?id=1620754.1620758>
- Binder, J. R., & Desai, R. H. (2011, November). The Neurobiology of Semantic Memory. *Trends in Cognitive Sciences*, 15(11), 527–536. doi:10.1016/j.tics.2011.10.001. pmid: 22001867
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral cortex*, 10(5), 512–528.
- Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., ... Weaver, K. E. (2011, January 15). Mapping Anterior Temporal Lobe Language Areas with fMRI: A Multi-Center Normative Study. *NeuroImage*, 54(2), 1465–1475. doi:10.1016/j.neuroimage.2010.09.048. pmid: 20884358
- Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016, December). Word meaning in the ventral visual path: A perceptual to conceptual gradient of semantic coding. *NeuroImage*, 143, 128–140. doi:10.1016/j.neuroimage.2016.08.068
- Bright, P., Moss, H., & Tyler, L. (2004, June). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and Language*, 89(3), 417–432. doi:10.1016/j.bandl.2004.01.010
- Bullinaria, J. A., & Levy, J. P. (2012). Extracting semantic representations from word co-occurrence statistics: Stop-lists, stemming, and SVD. *Behavior research methods*, 44(3), 890–907.
- Burgess, C., & Lund, K. (1995). Hyperspace analogue to language (HAL): A general model of semantic memory. In *Annual meeting of the Psychonomic Society, Los Angeles*.
- Chao, L. L. [Linda L.], Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature neuroscience*, 2(10), 913.
- Collins, A. M., & Quillian, M. R. (1969, April 1). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247. doi:10.1016/S0022-5371(69)80069-1
- Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, 126(5), 1193–1201.

- Crutch, S. J., & Warrington, E. K. (2004). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, 128(3), 615–627.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380(6574), 499.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, 92(1-2), 179–229.
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, 23(8), 3423–3431.
- de la Clergerie, E. (n.d.). DepGlove small server. Retrieved March 7, 2019, from <http://alpage.inria.fr/dep gql/proc ess>
- De Saussure, F. (1969). *Cours de linguistique generale: 3e ed.* Payot.
- de Saint-Exupéry, A. [A.]. (1943). *Le petit prince..*
- de Saint-Exupéry, A. [Antoine], & Wilkinson, D. (2011). *The little prince : a French/English bilingual book.* Rutland [England]: Omilia Languages Ltd.
- Deerwester, S. C., Dumais, S. T., Furnas, G. W., Harshman, R. A., CA, Landauer, T. K., ... Streeter, L. A. (1989, June 13). *United States Patent: 4839853 - Computer information retrieval using latent semantic structure.* 4839853. Bell Communications Research, Inc. Retrieved May 10, 2019, from <http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fs rchnum.htm&r=1&f=G&l=50&s1=4839853.PN.&OS=PN/4839853&RS=PN/4839853>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018, October 10). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv: 1810.04805 [cs]. Retrieved May 27, 2019, from <http://arxiv.org/abs/1810.04805>
- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., ... Tyler, L. K. (2000, June 1). Susceptibility-Induced Loss of Signal: Comparing PET and fMRI on a Semantic Task. *NeuroImage*, 11(6), 589–600. doi:10.1006/nimg.2000.0595
- Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005, May). Emotional and Temporal Aspects of Situation Model Processing during Text Comprehension: An Event-Related fMRI Study. *Journal of Cognitive Neuroscience*, 17(5), 724–739. doi:10.1162/0898929053747658
- Frank, S. L., & Willems, R. M. (2017, October 21). Word predictability and semantic similarity show distinct patterns of brain activity during language comprehension. *Language, Cognition and Neuroscience*, 32(9), 1192–1203. doi:10.1080/23273798.2017.1323109
- Freitas, A., Barzegar, S., Sales, J. E., Handschuh, S., & Davis, B. (2016). Semantic relatedness for all (languages): A comparative analysis of multilingual semantic relatedness using machine translation. In *European Knowledge Acquisition Workshop* (pp. 212–222). Springer.
- Fuster, J. M. (2004). Upper processing stages of the perceptionaction cycle. *Trends in cognitive sciences*, 8(4), 143–145.
- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience*, 26(18), 4917–4921.
- Goldstein, K. (1948). Language and language disturbances; aphasic symptom complexes and their significance for medicine and theory of language.
- Goldstein, K. (1971). The problem of the meaning of words based upon observation of aphasic patients. In *Selected Papers/Ausgewählte Schriften* (pp. 345–359). Springer.

- Gorno-Tempini, M. L. [M. L.], & Price, C. J. [C. J.]. (2001). Identification of famous faces and buildings: A functional neuroimaging study of semantically unique items. *Brain*, 124(10), 2087–2097.
- Gorno-Tempini, M. L. [Maria Luisa], Price, C. J., Josephs, O., Vandenberghe, R., Cappa, S. F., Kapur, N., ... Tempini, M. L. (1998). The neural systems sustaining face and proper-name processing. *Brain: a journal of neurology*, 121(11), 2103–2118.
- GramSchmidt Process. (2019, May 27), In Wikipedia. Page Version ID: 899073496. Retrieved May 29, 2019, from https://en.wikipedia.org/w/index.php?title=Gram%E2%80%93Schmidt_process&oldid=899073496
- Grossman, M., Koenig, P., Glosser, G., DeVita, C., Moore, P., Rhee, J., ... Gee, J. (2003, February 1). Neural basis for semantic memory difficulty in Alzheimers disease: An fMRI study. *Brain*, 126(2), 292–311. doi:10.1093/brain/awg027
- Harris, Z. S. (1954, August). Distributional Structure. *WORD*, 10(2-3), 146–162. doi:10.1080/00437956.1954.11659520
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301–307.
- Head, H. (1920). *Aphasia and kindred disorders of speech*. Cambridge University Press.
- Hill, F., Reichart, R., & Korhonen, A. (2015, September 25). SimLex-999: Evaluating Semantic Models With (Genuine) Similarity Estimation. *Computational Linguistics*, 41(4), 665–695. doi:10.1162/COLI_a_00237
- Hua, J., Xiong, Z., Lowey, J., Suh, E., & Dougherty, E. R. (2005, April 15). Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8), 1509–1515. doi:10.1093/bioinformatics/bti171
- Hughlings Jackson, J. (1879). On affections of speech from disease of the brain. *Brain*, 2, 203–222.
- Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016, April 27). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458. doi:10.1038/nature17637. pmid: 27121839
- Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012, December 20). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, 76(6), 1210–1224. doi:10.1016/j.neuron.2012.10.014. pmid: 23259955
- Jakobson, R., & Halle, M. (1963). *Fundamentals of Language*. S-Gravenhage: Mouton & Co.
- Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage*, 60(3), 1759–1770.
- Kutas, M., & Hillyard, S. A. (1984, January). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161. doi:10.1038/307161a0
- Lapesa, G., Evert, S., & Schulte im Walde, S. (2014). Contrasting Syntagmatic and Paradigmatic Relations: Insights from Distributional Semantic Models. In *Proceedings of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014)* (pp. 160–170). Proceedings of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014). doi:10.3115/v1/S14-1020
- Levy, O., & Goldberg, Y. (2014, June). Dependency-Based Word Embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 302–308). Baltimore, Maryland: Association for Computational Linguistics. Retrieved November 27, 2018, from <http://www.aclweb.org/anthology/P14-2050>

- Luria, A. R. (1976). *The working brain: An introduction to neuropsychology*. USA: Basic Books.
- Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in cognitive sciences*, 15(3), 97–103.
- Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information, henry holt and co. Inc., New York, NY, 2(4.2).
- Martin, A., & Chao, L. L. [Linda L]. (2001, April 1). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11(2), 194–201. doi:10.1016/S0959-4388(00)00196-3
- McRae, K., & Jones, M. (2013, March 11). *Semantic Memory*. doi:10.1093/oxfordhb/9780195376746.013.0014
- Merkle, E. C., You, D., & Preacher, K. J. (2016). Testing nonnested structural equation models. *Psychological Methods*, 21(2), 151–163. doi:10.1037/met0000038
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, January 16). Efficient Estimation of Word Representations in Vector Space. arXiv: 1301.3781 [cs]. Retrieved May 28, 2019, from <http://arxiv.org/abs/1301.3781>
- Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39–41.
- Miller, G. A. (1998). *WordNet: An electronic lexical database*. MIT press.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *science*, 320(5880), 1191–1195.
- Mollo, G., Cornelissen, P. L., Millman, R. E., Ellis, A. W., & Jefferies, E. (2017, January 11). Oscillatory Dynamics Supporting Semantic Cognition: MEG Evidence for the Contribution of the Anterior Temporal Lobe Hub and Modality-Specific Spokes. *PLOS ONE*, 12(1), e0169269. doi:10.1371/journal.pone.0169269
- Moseley, R. L., & Pulvermüller, F. (2014, May). Nouns, verbs, objects, actions, and abstractions: Local fMRI activity indexes semantics, not lexical categories. *Brain and Language*, 132, 28–42. doi:10.1016/j.bandl.2014.03.001
- Mummery Catherine J., Patterson Karalyn, Hodges John R., & Wise Richard J. S. (1996, August 22). Generating tiger as an animal name or a word beginning with T: Differences in brain activation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1373), 989–995. doi:10.1098/rspb.1996.0146
- Mummery, C., Shallice, T., & Price, C. (1999, May). Dual-Process Model in Semantic Priming: A Functional Imaging Perspective. *NeuroImage*, 9(5), 516–525. doi:10.1006/nimg.1999.0434
- Nakamura, K. [K.J.], Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., ... Schormann, T. (2000). Functional delineation of the human occipito-temporal areas related to face and scene processing: A PET study. *Brain*, 123(9), 1903–1912.
- Nakamura, K. [Katsuki], Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., ... Kojima, S. (2001, January). Neural substrates for recognition of familiar voices: A PET study. *Neuropsychologia*, 39(10), 1047–1054. doi:10.1016/S0028-3932(01)00037-9
- Noppeney, U., & Price, C. (2002, April). Retrieval of Visual, Auditory, and Abstract Semantics. *NeuroImage*, 15(4), 917–926. doi:10.1006/nimg.2001.1016
- Oldfield, R. C. (1971, March 1). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. doi:10.1016/0028-3932(71)90067-4

- Paivio, A. (2008). Mind and its evolution: A dual coding theoretical approach. *Psychol Rec*, 59, 295–300.
- Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011, January 11). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 201018711. doi:10.1073/pnas.1018711108. pmid: 21224415
- Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000, April). A Common Language Network for Comprehension and Production: A Contribution to the Definition of Language Epicenters with PET. *NeuroImage*, 11(4), 347–357. doi:10.1006/nimg.2000.0546
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007, December). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987. doi:10.1038/nrn2277
- Pecher, D., & Zwaan, R. A. (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge University Press.
- Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543).
- Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., ... Fedorenko, E. (2018, December 6). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(1), 963. doi:10.1038/s41467-018-03068-4. pmid: 29511192
- Ploux, S., & Ji, H. (2003, June 1). A Model for Matching Semantic Maps between Languages (French/English, English/French). *Computational Linguistics*, 29(2), 155–178. doi:10.1162/089120103322145298
- Pradet, Q., De Chalendar, G., & Baguenier-Desormeaux, J. (2014). Wonef, an improved, expanded and evaluated automatic french translation of wordnet. In *Proceedings of the Seventh Global Wordnet Conference* (pp. 32–39).
- Price, C. J. [Cathy J.], Devlin, J. T., Moore, C. J., Morton, C., & Laird, A. R. (2005). Meta-analyses of object naming: Effect of baseline. *Human Brain Mapping*, 25(1), 70–82. doi:10.1002/hbm.20132
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017, January). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. doi:10.1038/nrn.2016.150
- Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006, September 1). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, 6(3), 201–213. doi:10.3758/CABN.6.3.201
- Rowtula, V., Oota, S., Gupta, M., & Surampudi, B. R. (2018, September 30). A Deep Autoencoder for Near-Perfect fMRI Encoding. Retrieved January 21, 2019, from <https://openreview.net/forum?id=HJfYl4s0tX>
- Rubenstein, H., & Goodenough, J. B. (1965, October). Contextual Correlates of Synonymy. *Commun. ACM*, 8(10), 627–633. doi:10.1145/365628.365657
- Saedi, C., Branco, A., António Rodrigues, J., & Silva, J. (2018, July). WordNet Embeddings. In *Proceedings of The Third Workshop on Representation Learning for NLP* (pp. 122–131). Melbourne, Australia: Association for Computational Linguistics. Retrieved November 28, 2018, from <http://www.aclweb.org/anthology/W18-3016>
- Sagot, B. (2010). The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French. In *7th international conference on Language Resources and Evaluation (LREC 2010)*.

- Sagot, B., & Fier, D. (2008). Building a free French wordnet from multilingual resources. In *OntoLex*.
- Saumier, D., & Chertkow, H. (2002, November). Semantic memory. *Current Neurology and Neuroscience Reports*, 2(6), 516–522. pmid: 12359106
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123(12), 2400–2406.
- Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speecha positron emission tomography study. *The Journal of the Acoustical Society of America*, 120(2), 1075–1083.
- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive neuropsychology*, 20(3-6), 451–486.
- Simons, J. S., Koutstaal, W., Prince, S., Wagner, A. D., & Schacter, D. L. (2003, July). Neural mechanisms of visual object priming: Evidence for perceptual and semantic distinctions in fusiform cortex. *NeuroImage*, 19(3), 613–626. doi:10.1016/S1053-8119(03)00096-X
- Small, D. M., Jones-Gotman, M., Zatorre, R. J., Petrides, M., & Evans, A. C. (1997, July 1). A Role for the Right Anterior Temporal Lobe in Taste Quality Recognition. *The Journal of Neuroscience*, 17(13), 5136–5142. doi:10.1523/JNEUROSCI.17-13-05136.1997
- Smith, E. E., Rips, L. J., & Shoben, E. J. (1974, January 1). Semantic Memory and Psychological Semantics. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 1–45). doi:10.1016/S0079-7421(08)60451-X
- Todorovic, S., & Pallier, C. (2018). *Analyses IRMf lors de l'écoute de texte naturel*.
- Tranel, D., Grabowski, T. J., Lyon, J., & Damasio, H. (2005, August). Naming the Same Entities from Visual or from Auditory Stimulation Engages Similar Regions of Left Inferotemporal Cortices. *Journal of Cognitive Neuroscience*, 17(8), 1293–1305. doi:10.1162/0898929055002508
- Tsukiura, T., Mochizuki-Kawai, H., & Fujii, T. (2006). Dissociable roles of the bilateral anterior temporal lobe in face- name associations: An event-related fMRI study. *Neuroimage*, 30(2), 617–626.
- Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, 1, 381–403.
- Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277(5324), 376–380.
- Verdier, A., Lakretz, Y., & Pallier, C. (2018). *Encodage d'activité neuronale à partir de réseaux LSTM*. Neurospin. Saclay.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002, May). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, 5(5), 491–499. doi:10.1038/nn839
- Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, 106(4), 859–878.