Properties, Structures and Operations

Studies on language processing in the brain using computational linguistics and naturalistic stimuli

---

Alessandro Lopopolo

The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond.

More information can be found at `www.mpi.nl/imprs`

ISBN: 978-94-92910-21-9

Cover design: Panayiotis Tembriotis

Printed by: Ipskamp printing

Layout: This thesis was typeset with LaTeX 2_ε. It is based on the Clean Thesis style, adapted by Maria Otworowska.

The research reported in this thesis was conducted at the Center for Language Studies of the Radboud University Nijmegen, the Netherlands and was funded by the Netherlands Organisation for Scientific Research (NWO) Gravitation Grant 024.001.006 to the Language in Interaction Consortium.

Properties, Structures and Operations

Studies on language processing in the brain using computational linguistics and naturalistic stimuli

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op

dinsdag 12 januari 2021
om 14.30 uur precies

door

Alessandro Lopopolo

geboren op 6 januari 1985
te Crema (Italië)

Promotor
Prof. dr. A.P.J. van den Bosch

Copromotoren
Dr. K.M. Petersson (MPI)
Dr. R.M. Willems

Manuscriptcommissie
Prof. dr. M.T.C. Ernestus (voorzitter)
Prof. dr. P. Hagoort
Dr. A.E. Martin (MPI)
Dr. W.H. Zuidema (Universiteit van Amsterdam)
Dr. M. Otten (Universiteit van Amsterdam)

Properties, Structures and Operations

Studies on language processing in the brain using computational linguistics and naturalistic stimuli

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans to be defended in public on

Tuesday, January 12, 2021
at 14.30 hours

by

Alessandro Lopopolo

Born on January 6, 1985
in Crema (Italy)

Supervisor
Prof. dr. A.P.J. van den Bosch

Co-supervisors
Dr. K.M. Petersson (MPI)
Dr. R.M. Willems

Doctoral Thesis Committee
Prof. dr. M.T.C. Ernestus (chair)
Prof. dr. P. Hagoort
Dr. A.E. Martin (MPI)
Dr. W.H. Zuidema (University of Amsterdam)
Dr. M. Otten (University of Amsterdam)

Contents

1

Introduction

The study of language processing in the brain requires taking into consideration several of the properties of the linguistic stimuli and the operations the brain ought to perform in the course of processing itself. In the course of this dissertation, computational linguistic models are used to study: the properties of the stimulus that the brain is sensitive to, the structures that the brain constructs in order to comprehend it, and the putative basic operations that make processing possible.

The primary goal of this dissertation is to investigate the information types, structures, and operations that make up the sophisticated machinery underpinning natural language processing. The starting intuition is that the brain does not support a simple sequential pipeline where one type of information is processed at a time, and one type of representation is computed after another, but rather that it is composed of a multitude of co-existing and potentially synchronous and interacting processes.

In this dissertation, these intuitions are investigated using naturalistic stimulus paradigm data and computational linguistic modeling to link brain and behavioral data to different types of sequential properties of the stimulus, structural analysis, and basic processing operations. The reason for taking such an approach originates from the acknowledgment that language is a highly complex, yet structured, phenomenon, matched by the equally high complexity of the neural correlates that support it.

Properties of the stimulus – in the context of this dissertation – consist of the information that the brain extracts from the sequence of sounds or graphic signs that make up the stimulus itself. In the course of the dissertation, these properties are analyzed at the level of sequences of phonemes, words, and grammatical classes. Besides being sensitive to these sequential properties, the brain is also likely to perform a structural analysis of the incoming linguistic material. The brain constructs a network of relations between the lexical units recognized from the input stimulus. In the context of this dissertation, by basic operations, I refer to the retrieval of lexical information from the mental lexicon and the integration of this information in larger representations in the course of the analysis of the incoming linguistic stimulus.

1.1 Levels of language processing in the brain

In this section, I will further elaborate on the different aspects of language processing in the brain that are the focus in the present dissertation (see Figure 1.1). First I will discuss the fact that sequential properties of the stimulus may be decomposed by the brain in several types of information encompassing at least the surface word sequences, the sequences of grammatical categories these words assume in their context, and their phonotactic structure, at least when presented as a speech signal. I will then introduce the concept of syntactic structural analysis, which the brain is supposed to undertake to reach a sentence-level interpretation of the input; and the fact that these structures can also be decomposed in different typologies. Finally, I will introduce a somewhat orthogonal topic with regards to the types of operations that are assumed to enable the processing of the stimulus and the computation of its sentence-level interpretation: lexical retrieval and lexical integration into context.

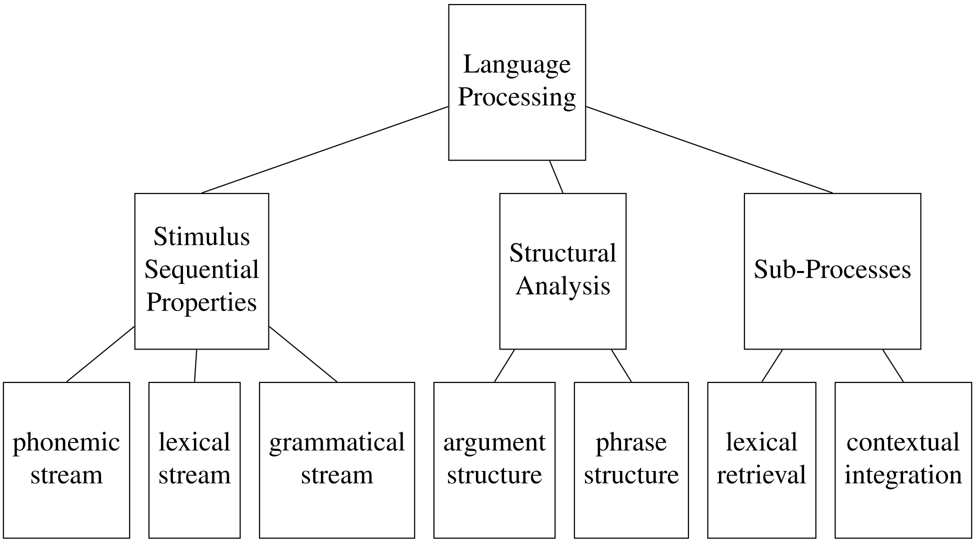


Figure 1.1: Conceptual map of the different aspects of language processing in the brain that will be considered and investigated using a combination of naturalistic stimuli, brain imaging data, and computational linguistics.

1.1.1 Properties of the sequential stimulus

(1) The man saw a brown dog in the park.

Sentence 1 is a fluent grammatical sentence in English. It consists of a string of symbols that, when spoken, unfold in time. When tokenized, it is composed of 9 word forms. Besides its surface form, Sentence 1 is composed by a sequence of 5 distinct grammatical categories (determiner, common noun, finite verb, adjective, and preposition)[1], and 17 distinct phonemes (Table 1.1).

The	man	saw	a	brown	dog	in	the	park.
w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
DT	NN	VBD	DT	JJ	NN	IN	DT	NN
[ðə]	[mæn]	[sɔː]	[ə]	[braʊn]	[dɑg]	[ɪn]	[ðə]	[pɑːk]

Table 1.1: Sentence 1 can be seen as a sequence of word forms, part of speech tags, and phonemes.

Although a natural language sentence is a linear sequence of surface forms, it allows being decomposed into different levels of information at the phrasal, word, and sub-word

---

[1] DT = determiner, NN = noun singular, VBD = verb past tense, JJ = adjective, IN = preposition. From Penn Treebank II Part-of-Speech tag set (`https://www.clips.uantwerpen.be/pages/mbsp-tags`).

level. These levels of information are what we might call, following traditional linguistic schools: syntax, lexico-semantics, and phonology (Jackendoff, 2003, 2007). The sentence as a **sequence of words** co-exists with the sentence as a **sequence of phonemes** and the sentence as a **sequence of grammatical categories**. Models of language processing typically decompose language into these co-existing levels of information (Hagoort, 2013a). These different levels are probably processed, at least partially, separately by the language faculty (Hagoort, 2005; Vigneau et al., 2006). It is, therefore, no surprise that the study of the neural basis of language comprehension has adhered to studying one of these types of information at a time. Kemmerer, 2014 offers a clear overview of the state of the art in each sub-discipline in the cognitive neuroscience of language.

In this dissertation, the neural bases of the three levels of information processing during language comprehension are investigated simultaneously, within one experiment. I want to know whether it is possible to fraction the neural signature of language processing of natural speech into different types of sequential information.

1.1.2 Structural analysis

Besides the multiple levels of sequential information mentioned above, language comprehension requires the construction of syntactic structures. From a neurobiological stance, sentential structural analysis is evidenced by the observation of the effects related to the syntactic structure of the stimulus on the patterns of brain activity involved in language processing. The neurolinguistic literature reports the existence of a network of mostly left-lateralized cortical regions whose activity appears to be modulated by the presence of linguistic stimuli that require, or allow, syntactic structural analysis. These areas usually include the left inferior frontal gyrus (IFG) (Caramazza & Zurif, 1976; Friederici et al., 2005; Tyler et al., 2008; Pallier et al., 2011; Zaccarella & Friederici, 2015; Zaccarella et al., 2015), the left posterior superior temporal gyrus (pSTG) and the left anterior temporal pole (ATP) (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2006, 2007; Baron et al., 2010; Baron & Osherson, 2011; Bemis & Pylkkänen, 2011, 2013; Westerlund et al., 2015; Bemis & Pylkkänen, 2013).

The involvement of several areas in structural analysis raises the question of whether these areas are sensitive to different types of structures or if the whole process is composed of a set of sub-processes. In this dissertation, I compare two types of syntactic structures represented by **phrase-structure grammar** (**PSG**) and **dependency grammar** (**DG**). These two grammars differ in several aspects. DG builds structures solely on the words and on binary relations holding between them. In contrast, PSG relies on grouping words in phrases that can, in turn, be grouped in larger phrases introducing a hierarchical structure composed by both surface forms (the words of the sentence) and non-observable abstract nodes that are assumed to be computed by the human brain. DG relates words through head-dependent relations, similar to predicate-argument structures. For this reason, it captures more closely the argument structure of the sentence. On the other hand, PSG captures hierarchical relations between phrasal groupings within the sentence.

In this dissertation, I will compare the predictive power of measures derived from these two types of grammars with regard to the activity in syntax-sensitive regions (left IFG, pSTG, ATP). These analyses are aimed to investigate the involvement of these areas in structural processing, and identify which areas are more sensitive to which type of structural description (PSG or DG). Data outside the neuroimaging domain can provide another way into the exploration of the type of structural analysis performed by humans during language processing. If the effects of structural analysis are detectable from patterns of brain activity, effects might also be noticed in behavioral or eyetracking data. In this dissertation, besides neuroimaging data, I also analyze patterns of eye movement during natural sentence reading, by testing the hypothesis that eye movements during reading reflect online syntactic analysis.

1.1.3 Basic operations

Cognitive and neural models of language comprehension typically distinguish two fundamental functional processes: **retrieval** and **integration** (Jackendoff, 2003; Hagoort & Levelt, 2009; Friederici, 2011; Kemmerer, 2014). First, the information relative to the incoming word is retrieved from memory. Memory refers to the long-term representation of word meaning from a putative "mental lexicon" (Damasio et al., 1996; Caramazza, 1996; Ullman, 2012). Second, an integration mechanism is needed that allows for these elements to be combined with each other in meaningful ways.

The distinction between retrieval and integration can be traced back to the observation that the language system seems able to deal with a virtually infinite number of utterances, which in turn seem to be composed by a limited, although flexible, set of primitives, such as phonemes, morphemes, words, and phrases. By storing these primitives in a hypothetical partition of long-term memory, recalling them when necessary, and combining them in a seemingly unbound number of configurations, humans can deal with a large variety of messages in a parsimonious, flexible and creative way. The present dissertation frames this distinction in terms of contextualization of lexical items, based on the assumption that the meaning of a word depends both on its lexical semantics and on the context in which it is embedded.

(2) In order to open a new account, you should go to a ***bank***.

(3) A fisherman is sitting with his rod on the ***bank*** of the river Thames.

Presented with Sentences 2 and 3, humans distinguish the meaning of *bank* as *"building or financial institution"* or as *"the shore of a river"* depending on whether it is encountered in the context of *"In order to open a new account, you should go to a **bank**"* or *"A fisherman is sitting with his rod on the **bank** of the river Thames"*. The string *"a new account"* steers the interpretation towards the financial domain, whereas the string *"a fisherman"* acts as bias towards a river-related interpretation of the word ***bank***.

It has been suggested that the human brain creates representations of words that are different according to such contextual cues (Willems & Casasanto, 2011). Brain activity related to language comprehension is composed of processes that involve different areas of the brain at different moments in time following the onset of the stimulus (Friederici et al., 2000; Friederici, 2002; Hagoort, 2005). It is, therefore, capital to show that the putative similarity between a model and a brain process regards not only areas associated with such process, but also that it does so in a time frame that is compatible with the time course of language processing. For this reason, I use a magnetoencephalographic (MEG) dataset collected during sentence reading. MEG records brain activity at the level of milliseconds, and with a reasonable anatomical resolution, making it ideal for a study interested in the when, and not only, the where of a specific neural process.

Binder et al., 2009 and Binder & Desai, 2011 provide a comprehensive picture of the cortical areas substantiating semantic memory. Memory, semantic memory in particular, is associated with the lateral portion of the left temporal cortex (middle temporal gyrus), parts of the inferior temporal gyrus, and the inferior parietal cortex. An important role is also hypothesized for the anterior portions of the temporal lobe (anterior temporal pole, ATP). The involvement of the ATP is confirmed by both studies on semantic dementia (SD) (Hodges et al., 1992; Rogers et al., 2004), and by a large neuroimaging literature (Tyler et al., 2004; Bright et al., 2005; Moss et al., 2004; Rogers et al., 2006). These findings were summarised by Patterson et al., 2007 and led to the formulation of the hub and spoke model, which posits that concepts are represented by a network of sensorimotor representations converging in the ATP which acts as a hub collecting and controlling modality-specific features to produce supra-modal representations. The input word lexical meaning is retrieved from semantic memory approximately between 150 and 300 ms after the onset of the word itself (Friederici, 2002; Humphries et al., 2007).

Integration operates on the representations retrieved from semantic memory, and in its most basic formulation, it consists of merging two linguistic tokens (e.g., two words) and creating a larger unit, such as a phrase or, more simply, a bi-gram. If one takes the concept of context in to consideration, integration can be seen as an operation that takes a token and embeds it in the context represented, for instance, by the other tokens making up the sentence in which it is presented. Brain imaging and brain lesion studies suggest that the inferior frontal gyrus, in interaction with areas in the perisylvian and temporal cortex, plays an essential role in lexical integration (Hagoort, 2005, 2013a). The anterior temporal areas have also been proposed to be involved in integration. For instance, several works have reported an increase in activity in these areas during the reading or listening of sentence as compared to word lists (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2006, 2007), under the assumption that the former condition requires word integration in larger syntactic units as compared to the later. This is confirmed by another series of studies which focused on simple phrasal processing, consisting of the composition of a wide range of phrasal and syntactic compositional types and cross-language and modality (Baron et al., 2010; Baron & Osherson, 2011; Bemis & Pylkkänen, 2011, 2013; Westerlund et al., 2015). As for its latency, integration is supposed to take place between 300 and 500 ms after stimulus onset (Berkum et al., 1999; Hagoort & van Berkum, 2007; Hagoort et al., 2009; Kutas & Federmeier, 2011).

1.2 Computational modelling

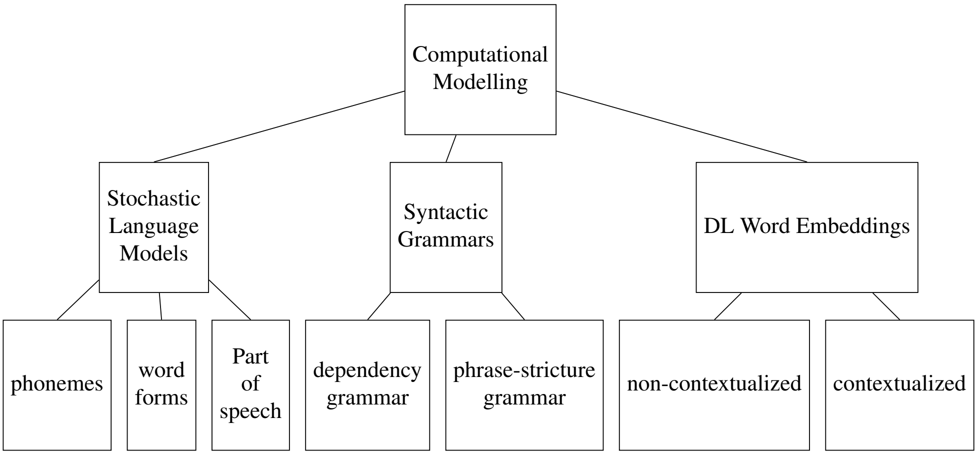


Figure 1.2: The different types of computational linguistic models used in this dissertation. The structure of this conceptual map reflects that of Figure 1.1.

Computational models are used in order to derive word, sub-word, or sentential quantitative descriptors of the stimuli presented to the participants of the experiment. The models fall under several different typologies and are all staples of the contemporary computational linguistic practice. The measures derived from such models are then employed as predictors or correlates of neural or behavioral activity elicited by the processing of the naturalistic stimuli described above (see Figure 1.2).

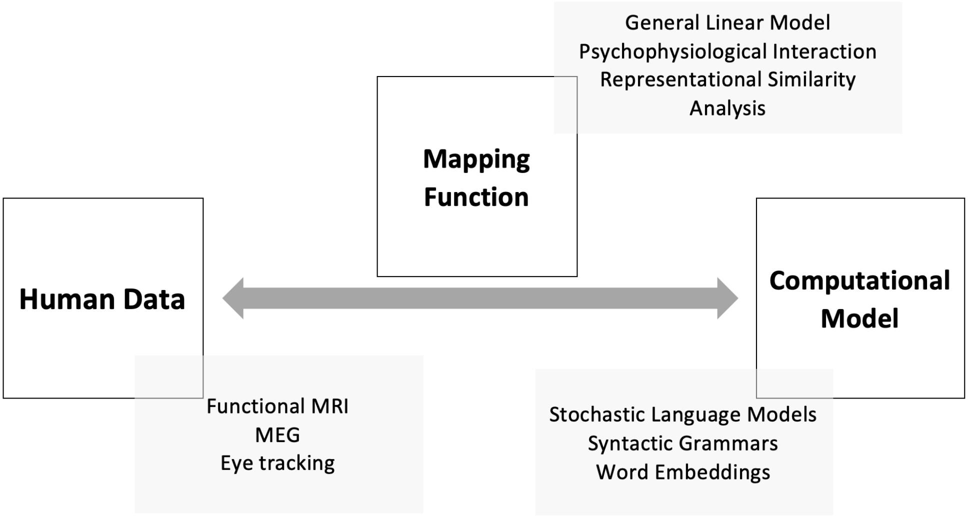


Figure 1.3: Schematics of the approach combining computational modeling and naturalistic stimuli to the answering of questions about natural language processing in the brain.

Figure 1.3 represents schematically the rationale behind using computational modeling to study language processing in the brain. Starting from a question concerning the type of properties or representations involved in language processing in the brain, a computational model is selected under the assumption that it might be a suitable approximation of those properties or representations. A computational model, therefore, acts as a computational hypothesis regarding an aspect of the neuro-cognitive phenomenon under analysis. Measures derived from computational modeling of the stimuli (introduced in this section) are mapped onto brain activity collected during the presentation of the same stimuli. This mapping can be achieved through several different *functions* that can consist in predicting brain activity given the computational measures (for instance using linear regression, like in a GLM) or by measuring the level of correlation between the way the brain and the model represent the stimulus space (using, for instance, representational similarity analysis, RSA presented by Kriegeskorte et al., 2008). The hypothesis that the model is a good approximation of the neural processes of language can then be quantified in several ways. For instance, in the case of an RSA, Pearson's correlation or other geometrical measures are used. In the case of predictive approaches, the prediction error or the analysis of the β-coefficients relative to the computational predictors are generally adopted.

The models presented here consist of stochastic language models, syntactic parsers, and word embeddings. Stochastic language models are used to model the sequential properties of the stimulus. Syntactic parsers are used instead to address questions about the structural analysis of the sentence that the brain is supposed to carry out. Word embeddings instead are adopted in order to study the types of operations and representations the brain employs during language processing. The following paragraphs will provide a schematic introduction of these models. More extensive descriptions will also be provided in further chapters.

1.2.1 Stochastic language models

In this dissertation, **language models** are implemented as n-order Markov models, also known as n-gram models, trained on a large collection of text. They are based on the simplifying assumption that the probability of word w_t depends on the previous two words only, that is, $P(w_t|w_1, ..., w_{t-1})$ is reduced to $P(w_t|w_{t-2}, w_{t-1})$.

For instance, in Sentence 1 the probability of encountering the word *saw* in the context provided by the sentence itself is given by $P(saw|the, man)$, which is the conditional probability of *saw* given the bi-gram the, man. In other words, a stochastic language model of Sentence 1 describes how likely it is to read (or listen to) the word form *saw* following words *the* and *man*. Similarly, indefinite article *a* is assigned conditional probability $P(a|man, saw)$. The probability of sentence initial tokens, such as *the* in this case, can be expressed in this framework by assuming the existence of place-holder symbols. Therefore the conditional probability of *The* is $P(The|START, START)$.

Descriptors of the sequential properties of the stimulus in accordance with a language model are then computed using stochastic measures such as **surprisal**, perplexity and entropy. Surprisal is computed as the negative logarithm of the conditional probability of w_t given w_{t-2}, w_{t-1}:

$$surprisal(w_t) = -\log P(w_t|w_{t-2}, w_{t-1})$$

If observing a word w has probability equals 1 (meaning that it is the only word that can be observed in that given moment), its surprisal will be 0. Conversely, the occurrence of a word that was considered almost impossible (i.e., with a probability close to 0) corresponds to a value of surprisal approximating infinite. In other words, surprisal can be thought of as the degree to which the perceived word w_t deviates from expectation: the less expected a word, the higher its surprisal. For the purpose of this dissertation, I will rely mainly on perplexity, consisting of an exponential transformation of the surprisal of encountering w_t given w_{t-2}, w_{t-1}.

1.2.2 Syntactic parsers

I distinguish between two main approaches to characterize the syntactic structure of a sentence: **phrase-structure grammars** (**PSG**), and **dependency grammars** (**DG**). Given a sentence, both grammars produce a hierarchical structure linking or grouping the words in a structure rooted in a governing node (the root node).

PSG defines parse structures of sentences as trees composed by terminal and non-terminal nodes. Non-terminal nodes correspond, usually, to phrasal categories as defined by the grammar in use, while terminal nodes – the leaf nodes of the tree – are assigned to the surface forms of the parsed sentence, i.e., its words. Phrase nodes are assigned labels corresponding to syntactic phrasal categories such as Noun phrase (NP), Verb phrase (VP), Adverbial phrase (AP), and Determiner phrase (DP).

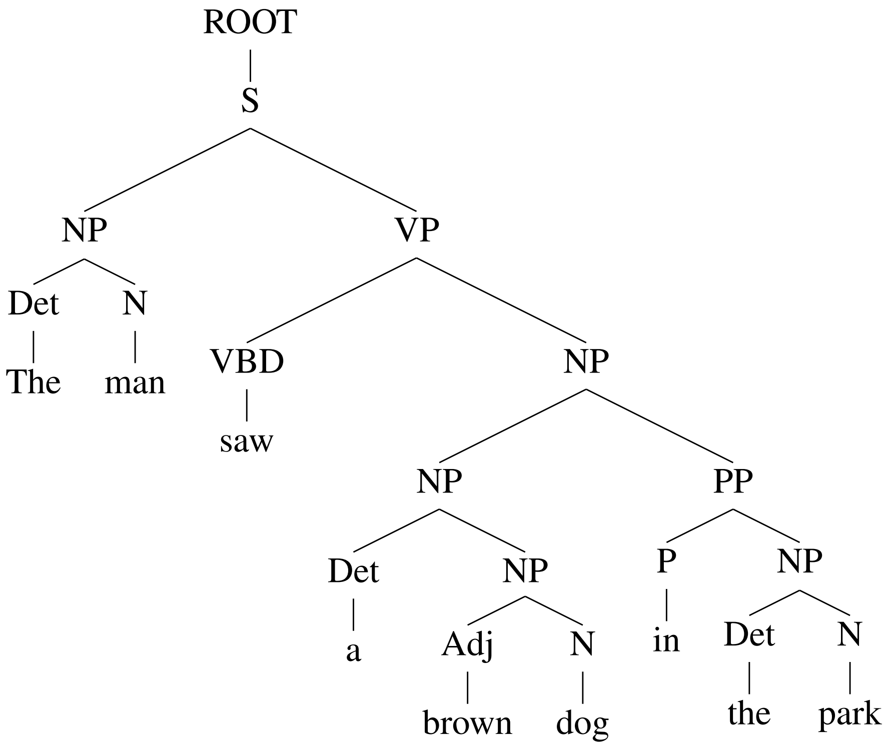


Figure 1.4: Phrase-structure parse of Sentence 1.

Figure 1.4 displays the parse of Sentence 1. The structure contains eight labeled phrase structures, including S, and constitutes a nested binary-branching tree. The words of the sentence (*the*, *man*, *saw*, *a*, *brown*, *dog*, *in*, *the*, and *park)* correspond to the terminal nodes. Following the structure of the parse tree in a top-down fashion: S branches into a NP and VP (noun and verb phrase, respectively). The left-hand child (NP) is composed of a determiner leaf node *the* and a noun *man*; whereas the right-hand child of S (VP) has in turn as left-hand child a terminal node (the finite verb *saw*) and as its right-hand side child another noun phrase (NP). This last NP branches off in another NP and in a prepositional phrase (PP). These two last phrases both split in a left-hand terminal child (respectively *a* and *in*) and in a NP as right-hand child. The latter two are both composed of terminal nodes (*brown*, *dog*, *the*, and *park*).

DG describes a sentence as a set of relations between pairs of words – a head and a dependent – composing it. The relations can be called dependencies and correspond to grammatical functions. The relations, and the words they link, are the only elements composing the structure (Tesnière et al., 2015; Mel'čuk, 1988; Nivre & Kübler, 2009). In a dependency structure, the finite verb is often taken to be the structural hub of the sentence. All other words are either directly or indirectly connected to the verb by dependencies.

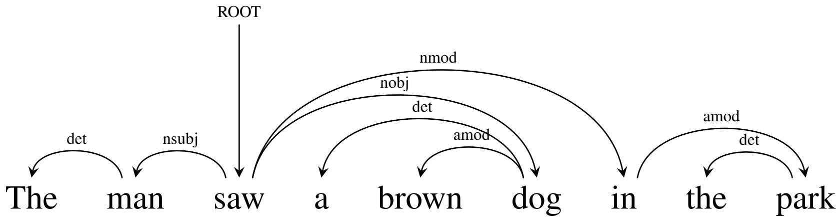


Figure 1.5: Dependency parse of Sentence 1.

Take for instance Sentence 1, the dependency graph in Figure 1.5 displays head-dependent relations holding the sentence together according to DG. As a way of example, the main verb (*saw*) acts as head for *man* and *dog*. It is worth noting how a dependent of one dependency relation can in turn be the head of another. For instance *dog*, one of the dependents of *saw*, is in turn head of *brown*, and of article *a*. Dependencies can be instantiated between words far apart in the sequential structure of the sentence.

It derives that relations holding between words are captured in structurally different manners by dependency parses and phrase-structure parses of the same sentence. A dependency structure directly captures the binary argument relations by means of a simple directed edge, whereas a phrase-structure relies on intervening higher order phrases.

1.2.3 Word embeddings

Linguistic units, such as words, can be represented by vectors of real numbers populating a high-dimensional space. This is achieved by distributional semantic models that leverage properties of the words and the contexts they appear in. For the purpose of this dissertation, I use two broad typologies of computational models developed for word representation generation: **non-contextualized models** and **contextualized models**.

Non-contextualized models generate representations that are independent from the context (sentence, paragraph, etc.) in which the represented word w is located. In this dissertation, I use word2vec (Mikolov et al., 2013) as model to generate non-contextualized word embeddings. It consists of a shallow, two-layer neural network. Once trained, it creates a high-dimensional vector space populated by word vectors, which are positioned in the space in such a way that words that share similar semantic and syntactic properties lay close to one another. Given Sentences 2 and 3 above, the model will return 27 embeddings (13 for each word in Sentence 2 and 14 for each word in Sentence 3). Of these 27 embeddings, the embedding of the word *bank* in Sentence 2 ($\vec{bank}^{w2v}_{@s1}$) and of *bank* in Sentence 3 ($\vec{bank}^{w2v}_{@s2}$) will be exactly the same, and the same would be the case for the four occurrences of the determiner *the*.

A contextualized model instead (Peters et al., 2018; Devlin et al., 2018; Melamud et al., 2016; McCann et al., 2017; Peters et al., 2017) assigns representations $\vec{w}$ that depend on the textual context in which the represented word w is located. The model used in this dissertation – ELMo (Peters et al., 2018) – is a recurrent bi-directional language model (biLM). The representation ELMo generates for a word consists of a concatenation of the activation of its two recurrent layers, and since the recurrent layer are designed to receives information about both the word w and the context it appears in, the representation (embedding) of word w is contextualized. For this reason, the embeddings that ELMo generates for *bank* in Sentence 2 ($\vec{bank}^{ELMo}_{@s1}$) will be different from the embedding of *bank* in Sentence 3 ($\vec{bank}^{ELMo}_{@s2}$), given the different contexts in which the two tokens of the same word are found.

1.3 Methodological notes

This dissertation aims to investigate the multiple levels that compose language processing from a sequential point of view, as well as from the point of view of the syntactic structures and of the types of operations involved. In order to do so, I have decided to adopt two innovative and complementary methodological paradigms: naturalistic stimuli and computational linguistic modeling. The naturalistic stimulus paradigm consists of the collection of data from participants that where presented with stimuli and conditions that resemble as much as possible real-life situations, such as narrative listening or reading, and with little or no experimental task forced onto them. This is in stark contrast with other approaches that rely on carefully constructed conditions and stimuli. Computational modeling – at least in the way it is meant here – refers to a series of tools adopted from computational linguistics and natural language processing (NLP), which are used to provide a detailed quantitative description of the stimulus.

Baggio & Hagoort, 2011 argue that the hypotheses underpinning theoretical models of the neuro-cognitive architecture of language processing are based on a *design stance* regarding how the brain must be organized and the dynamics that allow for it. They point out that it is difficult to test this hypothesis-derived architecture by using only task-based experiments such as vowel discrimination or lexical decision because such artificial tasks cannot be directly related to components of the cognitive architecture of the language system. In this dissertation, I put forward the idea that computational linguistic modeling applied to naturalistic stimuli allows making hypotheses about the architecture of the language system more testable. The main methodological contribution of computational modeling relies on the fact that it allows avoiding the limitations of task-oriented studies by exploiting the richness of naturally occurring sentences and by relying on a flashed out model of the process I want to investigate. In other words, computational linguistic modeling provides a more direct implementation of the process and not on the assumption that the process can be decomposed in orthogonal sub-processes that can be controlled by specific experimental tasks.

1.3.1 Limits of task-oriented paradigms

The limits of a traditional task-oriented approach to the study of language processing in the brain might be better understood through an example. Let us suppose that someone wants to answer the question of whether a particular region of interest in the brain (ROI) is sensitive to the phonemic structure characteristic of human languages. Traditional non-naturalistic paradigms might approach the question by presenting subjects with carefully constructed batteries of stimuli falling under different a priori defined conditions. In one condition, subjects might be presented with sequences of phonemes having properties matching those of their native language. Then they might also be presented with baseline or control conditions consisting in the presentation of random sequences of phonemes explicitly violating these properties and with white or pink noise sounds. Then the question is answered by comparing the ROI activity between these conditions. For instance, if the activity recorded during the structured phonemic sequences condition significantly diverges from activity recorded during the baseline conditions (e.g., noise or unstructured sequences), then the experimenter would tend to give a positive answer to the research question.

Such an approach has evident limitations. For instance, given the unfamiliarity with the tasks (few of us are expected, for instance, to listen to random lists of phonemes or to noise daily) subjects in this sort of experiment are often asked to attend to an ancillary task whose sole purpose is to ensure their attention.

Moreover, the problem with this kind of approach is that it is grounded on the assumption that language processing can be decomposed into isolated sub-processes and the entailed assumption that decomposing is possible at the level of the task or stimulation. This requires the assumption that the task and stimuli, carefully constructed to tap into a specific process, do not influence other processes, and that other processes do not indirectly influence the target process under examination. Similarly, it requires the assumption that the observed activity or behavior is the result of the linguistic conditions or task intended to elicit them, and not a by-product of the ancillary task adopted to ensure attention to such unfamiliar stimuli.

Moreover, this type of paradigm requires a leap of faith consisting of believing that language use (processing and production) under normal circumstances and language use under these restrictive experimental conditions involves the same processes supported by the same cognitive machinery and implemented by the same neural substrate.

1.3.2 Naturalistic stimuli

The rationale behind the use of naturalistic stimuli is the intuition that different sub-processes, types of information, and structures may be more directly studied by presenting the participants with a linguistic task as close as possible to their everyday experience, and by modelling the aspects of language processing one is interested in using current computational linguistic formalisms.

Unlike a task-oriented paradigm, naturalistic stimulus paradigms are not designed to elicit neural responses of a particular sub-process (and suppress or control for others) of language processing. On the contrary, the goal, or at least the accepted outcome, is the engagement of the whole neural machinery presumably responsible for language processing.

To put it in another way, since the stimuli are selected to be as close as possible to a real-life processing situation (reading, listening to a narrative), the elicited activity of the language processing system in the brain should be as similar as possible to the activity during the real-life homolog of the naturalistic stimulus itself. If the target is to collect and observe brain activity correlated to processing of language, naturalistic stimuli should provide a type of data that, although less controlled, is more ecologically valid than the one collected using task-oriented experimental paradigms.

The approach is in stark contrast to the task-oriented paradigm. The burden of the distinction between target processes in the task-oriented approach is carried by the tasks, stimuli, and stimulation paradigm in general. On the other hand, the task of distinguishing the effects of the subprocess(es) in the naturalistic-paradigm cannot be called upon the task, since, as I mentioned above, the task, or rather the lack of one, is meant to elicit all sub-processes involved in language understanding. In order, therefore, to tackle a specific process, naturalistic paradigms have to rely on the modeling of the stimulus itself. In the studies presented in this dissertation, modeling takes the form of several computational linguistics instruments, and it is used to generate quantitative descriptors of the stimulus, quantitative descriptors, which are then used as predictors or correlate of brain activity.

The descriptive powers of the measures compensate for the apparent uncontrolled nature of the naturalistic data. In other words, the burden of disentangling different components of a processing phenomenon is not posed on tasks or conditions forced onto the participants, but on the descriptors modeling the stimulus.

1.4 Research question(s)

Language processing is a complex phenomenon that requires the involvement of distinct sub-processes implemented by the human brain. In this dissertation I investigate the nature of these sub-processes, and, more specifically, whether they are implemented by distinct brain areas and signals. I analyse sequential, syntactic and operational sub-processes, using naturalistic language stimuli and computational modeling.

More specifically, with regard to the **sequential properties of the stimulus**, I intend to address the questions of whether the brain is sensitive to the stochastic properties of the stimuli captured by perplexity estimated by stochastic language models. Moreover, in line with the sub-processes question above, I am interested in assessing whether the sequential stochastic properties of the phonemes, words, and grammatical categories making up the stimulus are processed by distinct areas within the brain language network. The alternative hypothesis is that there is a single brain system responsible for processing the sequential stochastic properties of the stimuli.

As for the question regarding **syntax**, the present dissertation aims to assess the set of cortical areas sensitive to structural processing. Moreover, the focus will also be on what type of grammar better describes these structures (e.g., DG vs. PSG), and whether different areas are sensitive to different types of grammars.

Finally, concerning the **sub-operations** supporting language processing, the question is whether lexical retrieval and integration can be formalized in terms of contextualization of linguistic representations (i.e., word embeddings). I want to asses whether contextualized word embeddings model integration processing, and conversely, whether non-contextualized word embeddings model lexical retrieval in the brain.

1.5 Structure of the dissertation

Central to this dissertation are both theoretical questions and methodological considerations exposed in the previous sections of these introductory notes. The remainder of the dissertation is composed of chapters derived from a series of independent studies, each of which centered around one of the theoretical questions concerning sequential processing, hierarchical analysis, and basic operations.

Chapter 2 introduces the naturalistic stimulus datasets that will be used in the subsequent chapters. These datasets include an fMRI, an eyetracking and an MEG dataset. The fMRI and eyetracking data consist of naturalistic stimulus paradigm data, whereas the MEG data were collected during sentence reading.

Chapter 3 deals with the sequential properties of the stimulus and how the brain processes different levels of information. The analysis is conducted using measures derived from stochastic language models as predictors of BOLD activity at the whole-brain level.

Moving on, Chapters 4 and 5 present studies aimed at investigating the structural analysis of the incoming stimulus. After focusing on the sequential nature of the stimulus in Chapter 3, these chapters answer the question of whether the brain computes syntactic representations of the sentences that go beyond the mere sequential nature of the stimulus. Chapter 4 uses the same fMRI data as Chapter 3 and adopts measures derived from two types of syntactic structures as predictors of brain activity, both at the whole-brain level and at the level of single regions of interest. This chapter aims to elucidate the difference between argument structure – approximated by dependency grammar – and phrase structure – approximated by phrase-structure grammar – as hypotheses regarding the nature of the structural analysis performed by the brain. Chapter 5 instead investigates the hypothesis that patterns of eye movements during reading reflect, to a certain extent, the structure of the sentences as described by dependency structure of the sentences composing the stimulus.

Chapter 6 tackles an aspect of language processing that is somehow orthogonal to the ones addressed in the previous chapters. Starting from the assumption that the processing of the stimulus proceeds from the lexical to the sentential level, the chapter focuses on two basic operations supposed to bridge these two aspects: lexical retrieval and contextual integration. In order to do so, I use MEG data and compare them, both over anatomical regions and time, with the computational representations of the incoming stimulus, both at the lexical and at the contextual level.

Chapter 7 concludes the dissertation with some consideration on both the theoretical and methodological relevance of the studies presented in this dissertation.

2

Naturalistic Stimuli Datasets: fMRI, Eyetracking, MEG

This chapter contains the description of the datasets used in the studies included in the present dissertation. They consist of a functional magnetic resonance imaging, a magnetoencephalography, and an eyetracking dataset. Both the fMRI and eyetracking datasets were collected while subjects were presented with narrative texts. The MEG data were instead collected while participants were asked to read complex sentences. These datasets were used following the methodological choice of adopting a combination of naturalistic stimuli paradigm and computational linguistic modeling, as explained in the previous chapter.

This chapter is partially based on:

Lopopolo, Alessandro, Stefan L. Frank, Antal van den Bosch, Anabel Nijhof and Roel M. Willems. 2018. The Narrative Brain Dataset (NBD), an fMRI Dataset for the Study of Natural Language Processing in the Brain. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation 2018*. LREC, Miyazaki, 2018.

The methodological stance – expressed in Chapter 1 – calls for the adoption of data collected during naturalistic language stimulation, i.e., during listening or reading complex, close to real life, sentences or, when possible, long articulated narrative texts. The rationale behind this choice is the intuition that the various processes involved in language comprehension might be directly investigated by modelling their neural or behavioral correlates under conditions as close as possible to the ones they are likely to encounter in their everyday life. In other words, as mentioned in Chapter 1, instead of controlling for these processes by imposing artificial tasks to the subjects, this paradigm lets them process language as they might do in their everyday life, and then instead controls for variability and confound factors by carefully modelling them introducing computational predictors.

In this chapter, I introduce and describe the three different datasets I used in the course of the studies included in the present dissertation. They consist of three different sources of data (functional magnetic resonance imaging, eyetracking, and magnetoencephalography), and all but one were collected during the presentation of narrative texts. The fMRI dataset was collected while subjects were asked to passively listen to three Dutch narrative texts, while the eyetracking dataset was collected while a different pool of participants was asked to read a different set of Dutch narrative texts. Finally, the MEG data were acquired while Dutch complex sentences were visually presented to the participants.

2.1 fMRI narrative brain dataset

2.1.1 Introduction

The present fMRI dataset was created by recording the brain activity of 24 native speakers of Dutch during passive listening to three Dutch narrative texts: excerpts from audiobooks. This task and these stimuli are intended to be as naturalistic as possible. This dataset has already been used in several neuroscientific studies combining computational linguistic models and brain imaging analysis techniques, as exemplified in Section 2.1.4.

2.1.2 Data collection and pre-processing

Participants

Twenty-four healthy, native speakers of Dutch (8 males; mean age 22.9 years, range 18-31) without psychiatric or neurological problems, with normal or corrected-to-normal vision, and without hearing problems took part in the experiment. All participants except one were right-handed. Ethical approval was obtained from the CMO Committee on Research Involving Human Subjects, Arnhem-Nijmegen, The Netherlands (protocol number 2001/095), in line with the Declaration of Helsinki.

Procedure

The experimental paradigm consisted of passively listening to the three narratives (see Section 2.1.3) and their reversed versions (for a total of six sessions) inside the MRI scanner. That amounted to six experimental runs, all collected in one single fMRI session on the same day. Each story and its reversed speech counterpart were presented following each other. Reversed speech versions of the stories were created with Audacity 2.03[1]. Half the participants started with a non-reversed stimulus and half with a reversed speech stimulus. Participants were instructed to listen to the materials attentively, which in practice is only possible for three narratives, and not for the reversed speech counterparts. There was a short break after each fragment.

Stimuli were presented with Presentation 16.2[2]. Auditory stimuli were presented through MR-compatible earphones. After the scanning session, participants were tested for their memory and comprehension of the stories.

Scanner Parameter

Images of blood-oxygenation level-dependent (BOLD) changes were acquired on a 3-T Siemens Magnetom Trio scanner (Erlangen, Germany) with a 32-channel head coil. Pillows and tape were used to minimize participants' head movement, and the earphones that were used for presenting the stories reduced scanner noise. Functional images were acquired using a fast T2-weighted 3D echo-planar imaging sequence (Poser et al., 2010), with high temporal resolution (time to repetition: 880 ms, time to echo: 28 ms, flip angle: 14, voxel size: 3.5 × 3.5 × 3.5 mm, 36 slices). High resolution (1 × 1 × 1.25 mm) structural (anatomical) images were acquired using a T1 sequence.

Pre-processing

Preprocessing was performed using SPM8[3] and Matlab 2010b[4]. The first four volumes were removed to control for T1 equilibration effects. Rigid body registration was used to realign images. Images were realigned to the first image within each run. The mean of the motion-corrected images was then brought into the same space as the individual participant's

---

[1] `http://www.audacityteam.org`
[2] `https://www.neurobs.com`
[3] `http://www.fil.ion.ucl.ac.uk/spm`
[4] `http://www.mathworks.nl`

anatomical scan. The anatomical and functional scans were spatially normalized to the standard MNI template, and functional images were re-sampled to 2 × 2 × 2 mm voxel sizes. Finally, an isotropic 8-mm full-width at half-maximum Gaussian kernel was used to spatially smooth the motion-corrected and normalized data.

2.1.3 Linguistic Data

Narrative text used as stimuli presented to the human subjects consisted of three excerpts from three distinct literary novels extracted from the Spoken Dutch Corpus, *Corpus Gesproken Nederlands* (CGN) (Oostdijk, 2000).[5]

The excerpts were spoken at a normal rate, in a quiet room, by female speakers (one speaker per story). Stimulus durations were: Narrative 1 (CGN file fn1005) 3:49 min, Narrative 2 (CGN file fn1100) 7:50 min, and Narrative 3 (CGN file fn1090) 7:48 min.

The text below is an excerpt of the third narrative text used in this dataset (from Alan Jakobsen's novel *De Stalker*). Commas had already been removed in the file. Note the naturalistic nature of the linguistic material, composed by sentences of various length and rich vocabulary.

```
[...]
De donkerblauwe auto gleed als een slang door de bocht.
Zijn felle koplampen sneden als scheermessen door de stille
lucht. Hij reed door de buurt op zoek naar eten zijn prooi
opsnuivend. Na een trap op het  gaspedaal klom de snelheidsmeter
van de krachtige motor in minder dan drie seconden van zestig
naar negentig kilometer per uur. Zijn honger naar snelheid was
groot. Gezond verstand werd vermorzeld door angst boosheid en
woede op zoek naar de dood.
[...]
```

Table 2.1 contains summary information about the three narratives, including number of words, mean, and range of word duration in milliseconds.

---

[5] Narrative 1: from Peper, R., *Dooi*, L.J. Veen, 1999; Narrative 2: from Van der Meer, V., *Eilandgasten*, Contact, 1999; Narrative 3: from Jakobsen, A., *De Stalker*, De Boekerij, 1999

	# Words	Mean (s.d.)	Range
Narrative 1	622	273 (181)	4-1174
Narrative 2	1291	252 (160)	31-949
Narrative 3	1131	274 (183)	40-1221

Table 2.1: Summary information of the three narrative texts used as stimuli.

2.1.4 Published analyses of the current dataset

The present fMRI data have already been analyzed in several studies, demonstrating that naturalistic linguistic tasks and fMRI can yield interesting and meaningful results. Willems et al., 2016 have shown that entropy and surprisal predict brain activity in different brain areas. Frank & Willems, 2017 demonstrated that predictive measures (surprisal) and semantic association measures can be distinguished with regard to brain area sensitivity. Similarly, PoS, lexical, and phonological stochastic measures divide the cortical language network into non-overlapping sub-networks (Lopopolo et al., 2017). Part of the data was used by Nijhof & Willems, 2015 to investigate how individuals differently employ neural networks important for understanding others' beliefs and intentions, and for sensorimotor simulation while processing narrative language.

2.2 Eye tracker narrative dataset

The eye tracker data used in this dissertation was originally collected for a study on mental stimulation during literary reading by Mak & Willems, 2018 at Radboud University, Nijmegen, the Netherlands. For more details on data acquisition and preprocessing, we refer to the original publication.

2.2.1 Participants and stimuli

Data were collected from 102 participants (82 females, mean age 23.27, range 18–40), all of whom were native speakers of Dutch, with normal or corrected-to-normal vision. All participants gave written informed consent in accordance with the Declaration of Helsinki.

Stimuli consisted of three published short stories in Dutch. Stories 1 and 2 were written by contemporary Dutch writers, and Story 3 was translated from American English to Dutch. Their lengths were 2143, 2659, and 2988 words, respectively, and they required around 10-15 minutes each to be read. Similarly for the example text in Section 2.1.3, the text below is a portion of one of the 3 narratives presented to the participants during the acquisition of this narrative eye tracker dataset (from Rob van Hessen's short story *De mensen die alles lieten bezorgen*, contained in the collection *Hier wonen ook mensen*, van Essen, 2014).

[...]

'Dat is goed,' zei ik, en ik liep achter ze aan. Ze waren jong, ergens achter in de twintig. 'Komen ze niet terug, dan?' vroeg ik. 'Nee, dat kan niet met de trappen,' zei de agente. 'Hoe is het met haar?' vroeg ik. 'Dat kunnen we niet zeggen.' Het was onduidelijk hoe ze dat bedoelde: dat ze het niet wisten, of dat het informatie was die onder een of ander beroepsgeheim viel.

[...]

2.2.2 Data acquisition and pre-processing

For eye-movement data collection, a monocular desktop-mounted EyeLink1000+ eyetracking system was used (500 Hz sampling rate). Head movements were minimized using a head stabilizer, ensuring that all participants were seated at 108 cm from the screen.

The stimuli were presented using SR Research Experiment Builder software (SR Research, Ottawa, Canada). The stories were divided into 30 sections each. The stories were presented in counterbalanced order. After data collection, participants were presented with a comprehension questionnaire.

All fixations were checked to make sure that they did not drift off and enter a different interest area. If correction of the drifts was not possible, individual sections were excluded. Data for at least one section was removed for 40 participants. For four participants, the number of excluded sections exceeded six, resulting in the exclusion of one story for these participants.

Eight participants answered more than one comprehension question incorrectly for one of the three stories (four times for Story 2 and four times for Story 3), resulting in the exclusion of the data for one story reading for eight participants.

The dataset contains a total of 582,807 words across all participants and narratives.

2.2.3 Published analyses of the current dataset

As mentioned above, this dataset was originally collected for a study on mental stimulation during literary reading by Mak & Willems, 2018. de Vries et al., 2018 used this data to test assumption regarding the Deliberate Metaphor Theory (Steen, 2008) More recently, Faber et al., 2020 used this dataset to investigate individual reading style during literary reading as a function of the rate and type of word skipped.

2.3 MEG: MOUS dataset

The magnetoencephalographic data used in the context of this dissertation belong to the MOUS dataset (Schoffelen et al., 2019) collected at the Donders Centre for Cognitive Neuroimaging in Nijmegen, The Netherlands. For more details on the acquisition procedure, stimuli, preprocessing, and source reconstruction techniques, we refer to the original paper and to Schoffelen et al., 2017.

2.3.1 MEG data acquisition and pre-processing

The data were collected with a 275 axial gradiometer system (CTF). The signals were digitized at a sampling frequency of 1200 Hz (the cutoff frequency of the analog anti-aliasing low pass filter was 300 Hz). Head position with regards to the sensors was determined by 3 coils attached to the participant's head. Electro-cardiogram and the horizontal and vertical electro-oculogram were measured by 3 bipolar Ag/AgCl electrode pairs.

Electrocardiogram artifacts were identified based on their topography and subtracted from the data. The data was segmented into trials corresponding to activity recorded from -183ms before word presentation to a variable time after word presentation, depending on word length. Trials that contained artifacts (Eye movements and muscle contractions and jump artifacts in the SQUIDs) were excluded from further analysis. Next, the power line interference was estimated and subtracted from the data. The data were down-sampled to a sampling frequency of 300 Hz.

Source reconstruction was obtained using a linearly constrained minimum variance beamformer (LCMV) (Van Veen et al., 1997), estimating a spatial filter at 8,196 locations of the subject-specific reconstructed midcortical surface. The dimensionality of the data was reduced by applying an atlas-based parcellation scheme based on the Conte69 atlas (191 parcels per hemisphere). After that spatial filters were concatenated across vertices comprising a parcel, the first two spatial components were selected for each parcel. For more details on this procedure we refer to Schoffelen et al., 2017.

2.3.2 Subjects

We used the data of 74 subjects belonging to the MEG section of the MOUS dataset (Schoffelen et al., 2019). All subjects were Dutch native speakers, who were asked to silently read 120 Dutch sentences presented on a screen word by word, containing a total of 1377 words. All sentences varied between 9 and 15 words in length.

+ Gisteren had de brede bodyguard die de filmster beschermde een vrije dag.
+ De sportarts die de hockeyer met de gescheurde pezen adviseert heeft er veel verstand van.
+ De journalist die de beroemde bondscoach interviewt wil veel weten.
+ De goochelaar die de vrijwilliger uit het enthousiaste publiek instrueert is onduidelijk.
+ De detective die criminelen opspoort krijgt een vette beloning van de staat.
+ Het derde getal dat de oplettende toehoorder signaleert is zes.
+ Onlangs gaf de jongeman die de populaire portier inhuurde een groot feest.
+ Het elfje dat de sterke beren betoverde was erg vriendelijk.
+ De schilder die de knappe prinses tekent zit onder de verf.
+ Nora die warme dekens voor arme mensen weeft is gelukkig.

Examples of sentences using during the acquisition of these data are presented in the text above, punctuation, except from the final full stop, was not presented.

2.3.3 Stimulation paradigm

The sentences were presented visually with an LCD projector, with a vertical refresh rate of 60Hz situated outside the MEG scanning room, and projected via mirrors onto the screen inside the measurement room. All stimuli were presented in a black mono-spaced font on a gray background at the center of the screen within a visual angle of 4 degrees. Sentences were presented word-by-word with a mean duration of 351ms for each word (minimum of 300ms and maximum of 1400ms, depending on word length). The duration of the stimuli was determined taking into account both the word length, the number of words in the whole sentence, and the number of letters within each word. Each word was separated by an empty screen for 300ms before the onset of the next word.

3

Sequential Processing of Language at Different Information Levels

Linguistic stimuli consist of sequences of speech sounds or graphemes. In this chapter, I will model how they are processed in the brain by using stochastic language models. Previous studies (Frank et al., 2015) have already shown that such models can be useful tools for studying how language is processed as a sequence of symbols unfolding in time. Nonetheless, most of these studies have focused only on sequences of words, not addressing the fact that language processing involves the simultaneous processing of information at the phonological, syntactic, and lexical levels. In this chapter, I will track these three distinct levels of information in the brain by using stochastic measures derived from stochastic language models to detect neural correlates of phoneme, part-of-speech, and word processing in an fMRI experiment. Brain activity time-locked to each word is predicted by the probabilistic perplexity derived from these three models. The results show that the brain keeps track of the statistical structure of lexical, syntactic, and phonological information in distinct areas.

This chapter is based on:

Lopopolo, Alessandro, Stefan L. Frank, Antal van den Bosch, and Roel M. Willems. 2017.Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain. *PLOS ONE* 12 : 1–18.

In Chapter 1, I have introduced the hypothesis that sequences of words making up the surface form of a sentence can also be analyzed as sequences of phonemes composing them and sequences of grammatical categories describing their combinatorial properties in the context of the sentence itself. These correspond to different levels of information at the phrasal, word, and sub-word level, and are traditionally referred to as: syntax, lexico-semantics, and phonology (Jackendoff, 2003, 2007). Models describing how the brain processes language typically take these levels into consideration (Hagoort, 2013a), and a common assumption is that they are probably processed by separate sub-components of the language faculty (Hagoort, 2005; Vigneau et al., 2006; Kemmerer, 2014).

This chapter investigates the neural bases of these three types of sequential information by using one fMRI dataset and predictors derived from stochastic language models trained on words, grammatical categories, and phonemes. The goal is to show that the processing of words, phonemes and grammatical categories sequences elicit distinct neural signatures within the language network in the brain.

3.1 Language processing as a sequential stochastic process

A number of studies (Bar, 2011; Bubic et al., 2010; K. Friston & Kiebel, 2009; Summerfield & Egner, 2009; Clark, 2013) have advanced the hypothesis that the brain employs predictive coding strategies in perception. The hypothesis is that after processing the first $t-1$ elements of a sequence of stimuli (i.e., $x_1, ..., x_{t-1}$), the human brain assigns a conditional probability $P(x_t|x_1, ..., x_{t-1})$ to each potential element x that can follow at time t. These expectations influence the way the actual observed x_t is processed eventually. Deviations from expectations are usually quantified in terms of surprisal or perplexity, which have been shown to explain both behavioral and neural correlates of perceptual and higher cognitive processing.

In the domain of language processing, word surprisal has been used to predict a wide range of behavioral correlates. It has been found to predict the duration of spoken words, with shorter words being used in less surprising situations (Piantadosi et al., 2011; Mahowald et al., 2013). Following Hale, 2001 and Levy, 2008, surprisal has been hypothesized to be proportional to the cognitive effort required to integrate a word into the current context. This has been confirmed by observing that it correlates with reading times (Frank & Thompson, 2012; Monsalve et al., 2012; Frank, 2013; Smith & Levy, 2013). Reading time has also been shown to be correlated to surprisal of the syntactic category (part-of-speech; PoS) of the word being read (Boston et al., 2008; Frank & Bod, 2011). Moreover, Monsalve et al., 2012 showed that PoS and word surprisal have independent effects on reading times. This behavioral result suggests that the PoS of words in sentential context is a valid representation of linguistic information relevant for processing and that computing probabilistic measures on them returns a model that has significant predictive power. These measures have also been successfully applied to the prediction of brain activity. It was found that the amplitude of the N400 event-related potential (ERP) component elicited by words in sentences correlates with word surprisal values (Frank et al., 2015; Parviz et al., 2011). The fact that surprisal correlates with the amplitude of a classical ERP component related to language comprehension (Kutas & Federmeier, 2000) is another source of evidence that stochastic language modeling is a neuro-cognitively valid approximation of sentence comprehension. In a recent paper, Willems et al., 2016 applied surprisal and entropy to an fMRI dataset to predict brain activity in different cortical and subcortical areas during naturalistic language comprehension. They observed that different areas differentially code for statistical stimulus properties by selectively correlating with one or the other measure.

3.2 Probabilistic streams

My starting point is twofold. On the one hand, my hypothesis is that different types of information correspond to different streams of processing implemented in separable networks in the brain. Several linguistic models, for instance, separate phonological, semantic, and syntactic processing in different neural loci or processing streams (Kemmerer, 2014). One way to operationalize this search for parallel streams is to model the processed linguistic input as composed by three parallel levels of representation corresponding to its phonological, lexical, and syntactic profile. On the other hand, following the findings exposed in section 3.1, the processing in these separate streams is modeled as sequential and incremental, and sensitive to the stochastic properties of the information it is applied to. The difference between different streams is that the probabilistic relations are computed not only on surface forms (bare words, so to speak) but on the phonemic transcription and the grammatical categories of the words in order to disentangle the different levels of representation.

The fact that language can be studied as a stochastic process does not necessarily mean that subcomponents of language correspond to distinct stochastic processes that are detectable in the brain. One scenario could be that the areas sensitive to stochastic properties of the input are the same, independently from the level of annotation on which such measures have been computed. Nastase et al., 2013 investigated whether there exist areas in the brain that are sensitive to probabilistic properties of the incoming signal, independently from its sensory modality, or if, conversely, sensitivity to such properties is an intrinsic characteristic of domain-specific areas. Their approach consisted of looking for areas coding for the degree of disorder - quantified by Markov entropy - in a temporally unfolding sensory input of two distinct modalities: auditory and visual. Their results show a modality-specific sensitivity to input entropy, implemented in modality-specific systems of sensory cortices (for visual stimuli: the early visual cortex, the anterior cingulate, and the intraparietal sulcus; for acoustic stimuli: inferior frontal, lateral temporal, and supplementary motor regions). Ventral premotor and central cingulate cortices were identified as possible candidates for modality-general uncertainty processing, exhibiting sensitivity to disorder in both modalities.

I have decided to approach the problem of disentangling phonology, lexical, and syntax by using language stimuli that are not explicitly designed to study one of these levels in isolation (Frank et al., 2015; Willems et al., 2016). Using one stochastic measure computed on three distinct levels of annotations of the same linguistic stimulus, I want to investigate first of all level-specific processing, what it may be referred to as streams of information. On the other hand, I am interested in investigating the issue of whether there exists a central supramodal stochastic processor of the brain (what is called modality-independent in Nastase et al., 2013) by finding areas that are sensitive to stochastic measures independently from the level of information they have been computed on.

3.3 Materials and methods

3.3.1 fMRI data

For the purpose of this study, I have used the fMRI narrative brain dataset described in Chapter 2 Section 2.1. The dataset consists of data of 24 native Dutch speakers collected while the participants were asked to listen to the spoken presentation of three Dutch narrative texts.

3.3.2 Estimation of stream-wise stochastic properties

The three levels of information – phonological, syntactic, and lexical – are distinguished by applying three different levels of annotation to the stimulus narratives. At the phonological level, the words in the running text are transformed into a sequence of phonemes. The syntactic level is approximated by the sequence of fine-grained syntactic categories corresponding to the words of the texts, also known as parts of speech (PoS). The lexical level consists of the sequence of surface lexical forms composing the texts. In the sections below, I describe the characteristics of each stream and their common computational properties.

The lexical level is the sequence of words constituting the sentences of the book fragments. At this level, a sentence can be rewritten as a sequence $w_1, \ldots, w_n$ of symbols w_i belonging to the vocabulary V containing all the word forms, as illustrated in Table 3.1 for the Dutch sentence *ze staat stil en kijkt een poosje naar een punt in de verte* (taken from one of the stimulus narratives used in the present study).

Ze	staat	stil	en	kijkt	een	poosje	naar	een	punt	in	de	verte
w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}	w_{13}

Table 3.1: The lexical stream is obtained from the simple sequence of word forms in the stimulus presented to the subjects.

The phonological stream can be defined as the sequence of phonemes composing each single word in the sentence. Therefore, the sentence in Table 3.1 can be rewritten as a sequence $p_1^1, p_1^2, ..., p_1^m, ..., p_{13}^1, ..., p_{13}^o$ where p_j^i refers to the i^{th} phoneme of the j^{th} word in the sentence. Table 3.2 contains the phonetic transcription of the example sentence already presented in Table 3.1.

Ze	staat	stil	en	kijkt	een	poosje	naar	een	punt	in	de	verte
[zə]	[stat]	[stɪl]	[ɛn]	[kɛɪkt]	[ən]	[poʃə]	[nar]	[ən]	[pynt]	[ɪn]	[də]	[vɛrtə]

Table 3.2: The phonological stream is obtained from the phonetic transcription of the words of the stimulus.

Finally, the words in the stimuli are assigned with their syntactic categories or part of speech tags (PoS). Parts of speech are a basic ingredient of most language technology systems and act as shallow (i.e., non-hierarchical) syntactic starting point for many other tasks, including semantic role assignment and dependency and constituent syntactic parsing. They usually consist of a basic set of grammatical categories such as nouns (N), verbs (WW, in the Dutch tags used here), modifiers and determiners. They capture, when considered in context, shallow, yet robust, combinatorial constraints that abstract away from the lexical information within the surface forms. The tagset employed here was the one employed by CGN (the corpus from which the stimuli for my experiments were taken) and comprises 320 tags (see Table 3.3). Besides 13 base tags, this method explicitly assigns morpho-syntactic sub-category features to the base tags containing information such as gender, number, form, and so on.

This tagset closely follows the practices of the Dutch Grammar ’Algemene Nederlandse Spraakkunst’ (ANS, Haeseryn et al., 1997). Table 3.4 contains an example of PoS annotation of the example sentence presented in Table 3.1 above.

Dutch POS	tag	#	English equivalent	Example
substantieven	N	18	Nouns	het kind
adjectieven	ADJ	30	Adjectives	de mooie huizen
werkwoorden	WW	21	Verbs	ik kom
telwoorden	TW	11	Quantifiers	vier cijfers
voornaamwoorden	VNW	188	Pronouns	ik
lidwoorden	LID	9	Articles	de hond
voorzetsels	VZ	3	Prepositions	in het hospitaal
voegwoorden	VG	2	Conjunctions	Jan en Peter
bijwoorden	BW	1	Adverbs	gisteren
tussenwerpsels	TSW	1	Interjections	hoera!
speciale tokens	SPEC	35	special forms	
leestekens	LET	1	Punctuation	.
TOTAL		320		

Table 3.3: Summary of the types of grammatical categories (POS) and the number of sub-categories used to approximate sequential syntactic information processing.

Ze	staat	stil	en	kijkt	een	poosje	naar	een	punt	in	de	verte
VNW	WW	ADJ	VG	WW	LID	N	VZ	LID	N	VZ	LID	N

Table 3.4: The same sentence from Tables 3.1 and 3.2 annotated with fine-grained grammatical information using the POS tags described above.

3.3.3 Computing stochastic measures

The conditional probabilities required for obtaining perplexity values for the lexical and PoS streams are estimated by a second-order Markov model, also known as trigram model, trained on a large collection of text. It is based on the simplifying assumption that the probability of word w_t depends on the previous two words only, that is, $P(w_t|w_1, ..., w_{t-1})$ is reduced to $P(w_t|w_{t-2}, w_{t-1})$. Surprisal is computed as the negative logarithm of the conditional probability of w_t given w_{t-2}, w_{t-1}:

$$surprisal(w_t) = -\log P(w_t|w_{t-2}, w_{t-1})$$

If the observed word's probability equals 1, observing it yields a surprisal of 0. Conversely, the occurrence of a word that was not among the words considered possible (i.e., has zero probability) corresponds to infinite surprisal. Surprisal can be thought of as the degree to which the actually perceived word w_t deviates from expectation. Perplexity, as adopted here, consists in an exponential transformation of the surprisal of encountering w_t given w_{t-2}, w_{t-1}.

$$ppl(w_t) = 2^{surprisal(w_t)} = 2^{-\log P(w_t|w_{t-1})}$$

The dataset from which probabilities $P(w_t|w_1, ..., w_{t-1})$ are estimated is a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the Dutch Corpus of Web (NLCOW2012 Schäfer & Bildhauer, 2012). For lexical perplexity, each word of the experimental texts is assigned a value computed by SRILM (Stolcke, 2002).

The PoS perplexity is computed analogously. Instead of using the surface forms of the training and stimulus set, the trigram model was trained on the PoS-tagged version of the same 10 million sentences subset of NLCOW2012. The tagging was performed using the Frog toolbox for natural language processing of Dutch text (Daelemans & van Den Bosch, 2005).

The phonological information was extracted from the phonemic transcription of each word in the stimulus set. I used a memory-based grapheme-phoneme converter (Busser et al., 1999) trained on CELEX 2 (Baayen et al., 1995). Once every word is transcribed as a sequence of phonemes, trigrams were extracted and conditional probabilities $P(p_t|p_{t-1}, p_{t-2})$ were computed using WOPR[1] trained on CELEX 2 (Baayen et al., 1995). Once phoneme-wise perplexity is computed, the phonetic perplexity of each word of the stimulus is computed as the average value across the phonemes of that word.

3.4 Data analysis

At the single-subject level, the observed BOLD time course in each voxel is subjected to a regression analysis, testing for voxels in which the covariates of interest (word, PoS, and phonological perplexity) explain a significant proportion of variance of that voxel's time course (K. Friston, 1995). Before the actual analysis, one regressor modeling the duration of every single word was created for each story. This regressor was convolved with the hemodynamic response function to account for the delay in BOLD activation respective to stimulus presentation. The word duration regressor and the covariates for a story were also fitted to the data of the reversed speech version of that story. This served as a control condition since the regressors and covariates are essentially meaningless for the reversed speech data.

---

[1] `https://ilk.uvt.nl/wopr/`

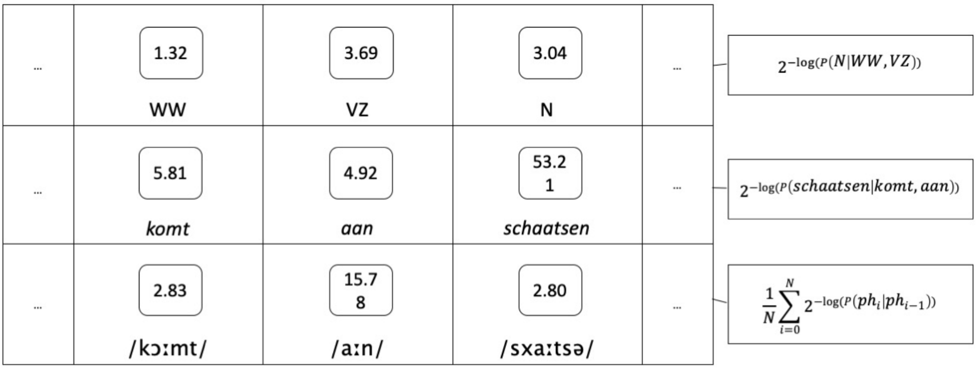


Figure 3.1: Example of annotation of the passage *"... komt aan schaatsen ..."* from one of the 3 narrative texts used during stimulation. The top row describes the PoS-level, the middle row the lexical level, and the bottom row the phonological one. The graph shows how the perplexity measures are computed for the last word (*schaatsen*), and of its PoS-tag (*N*) as a function of the probability of encountering it in that particular context. The phonological measures is instead the average perplexity of the phonemes composing a word (/sxaːtsə/).

Three covariates were computed containing each word's word, PoS, and phonemic perplexity measures, constituting the regressors of interest modeling the three information streams introduced above. Besides these, log2-transformed lexical frequency per word was computed using the Subtlex NL corpus (Keuleers et al., 2010), log2-transformed PoS frequency per word was computed using the CGN corpus (Oostdijk, 2000), and log2-transformed phoneme frequency average per word was computed using CELEX 2 (Baayen et al., 1995). They were used as regressors of no interest to statistically factor out effects of word, PoS, and phoneme frequency. The estimates from the motion correction algorithm (three rotations and three translations per run) were additionally added as regressors of no interest.

The modeled time courses from all six runs (three stories and three reversed speech stimuli) were combined in one regression model, with separate constant terms per run, but the same regressors for real and reversed speech. The analyses were conducted at the whole-brain level. The difference in the effect of the regressor of interest between the real and reversed speech sessions was used as input to the group-level statistics. Statistical differences were assessed by computing the t-statistic over participants of this difference score (real vs. reversed speech) for each voxel in the brain. The resulting multiple comparisons problem was solved by means of combining a $P < 0.005$ voxel threshold with a cluster extent threshold determined by means of 1,000 Monte Carlo simulations, after estimation of the smoothness of the data applied for each separate contrast. The combination of a voxel-level threshold with a cluster extend threshold is a good compromise between statistical sensitivity on the one hand and false positive error control, on the other hand (Bennett et al., 2009; Woo et al., 2014). The simulations took the amount of autocorrelation in the data into account, as suggested in the literature (Bennett et al., 2009; Woo et al., 2014). The scripts used were taken from[2]. All clusters of size display results significant at the $P < 0.05$ level, corrected for multiple comparisons.

3.4.1 Relation between the regressors

The aim of this study is to assess whether different types of linguistic information can be traced in the brain and if this can be achieved by using stochastic measures of perplexity in line with the predictive brain hypothesis. In order to assess whether word, PoS, and phoneme perplexity capture different kinds of information, I conducted a preliminary analysis consisting in computing their pairwise correlations. Table 3.5 reports these correlations (Pearson's r). Both 3-gram perplexity (ppl) and 1-gram frequency (freq) computed at lexical, PoS, and phonological levels are included.

	Lex_ppl	**PoS_ppl**	**Pho_ppl**	**Lex_freq**	**PoS_freq**	**Pho_freq**
Lex_ppl	1	0.046	0.011	−0.466	−0.092	0.080
PoS_ppl		1	−0.012	−0.015	−0.491	0.000
Pho_ppl			1	−0.016	−0.000	−0.017
Lex_freq				1	0.070	−0.060
PoS_freq					1	0.105
Pho_freq						1

Table 3.5: Correlation between the stochastic measures used in the analyses.

The correlations between perplexity measures reported in Table 3.5 are fairly low, even between lexical and PoS perplexity (0.046). These results indicate that there is no strong relation between the regressors I have employed in these fMRI analyses and that they may capture different types of information. Correlation between lexical perplexity and frequency is -0.466, and correlation between PoS perplexity and frequency is -0.491. These negative correlations between perplexity and frequency measures are predictable: the less frequent an item is, the higher is the overall perplexity of encountering it.

---

[2] `https://www2.bc.edu/~slotnics/scripts.html`

3.5 Results

In this section, I present the results of the analyses conducted using the three perplexity measures as regressors of interest.

3.5.1 Lexical stream

Table 3.6 lists the areas that show significant activity with regard to the word-based perplexity regressor. This network is displayed in Fig 3.2 and it encompasses large portions of the left inferior temporal gyrus (l-ITG), including the fusiform gyrus (l-FG). Both left and right posterior banks of the superior temporal gyrus (rl-STG) are part of this network, together with parts of the left anterior superior temporal gyrus.

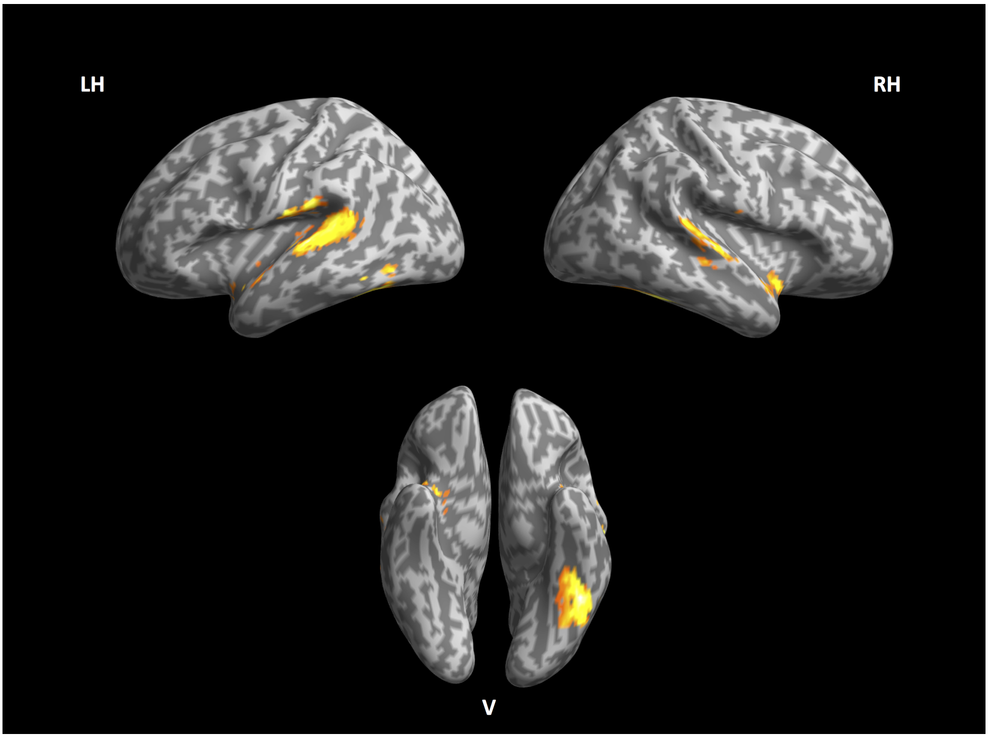


Figure 3.2: Lexical stream. Left, right and ventral view of inflated cortex plot of the lexical stream.

Region	MNI	size	t-value max
left inferior temporal gyrus - fusiform gyrus	-44 -48 -14	924	5.92
left posterior superior temporal gyrus	-56 -26 8	1876	6.29
left middle temporal gyrus	-58 -22 2		4.59
left anterior superior temporal gyrus (TP)	-40 2 -16	121	5.72
right posterior superior temporal gyrus & sulcus	64 -10 -2	1436	4.72

Table 3.6: Significant effects of lexical stream perplexity.

3.5.2 Syntactic stream

Fig 3.3 shows the cortical network corresponding to PoS-based perplexity. These include the left middle temporal gyrus and sulcus (l-MTG and l-MTS) and right middle temporal sulcus (r-MTS). The bilateral precentral sulcus is also activated. Large portions of the superior frontal gyrus are also sensitive to this regressor. The list of areas with the coordinates of their activation peaks can be found in Table 3.7.

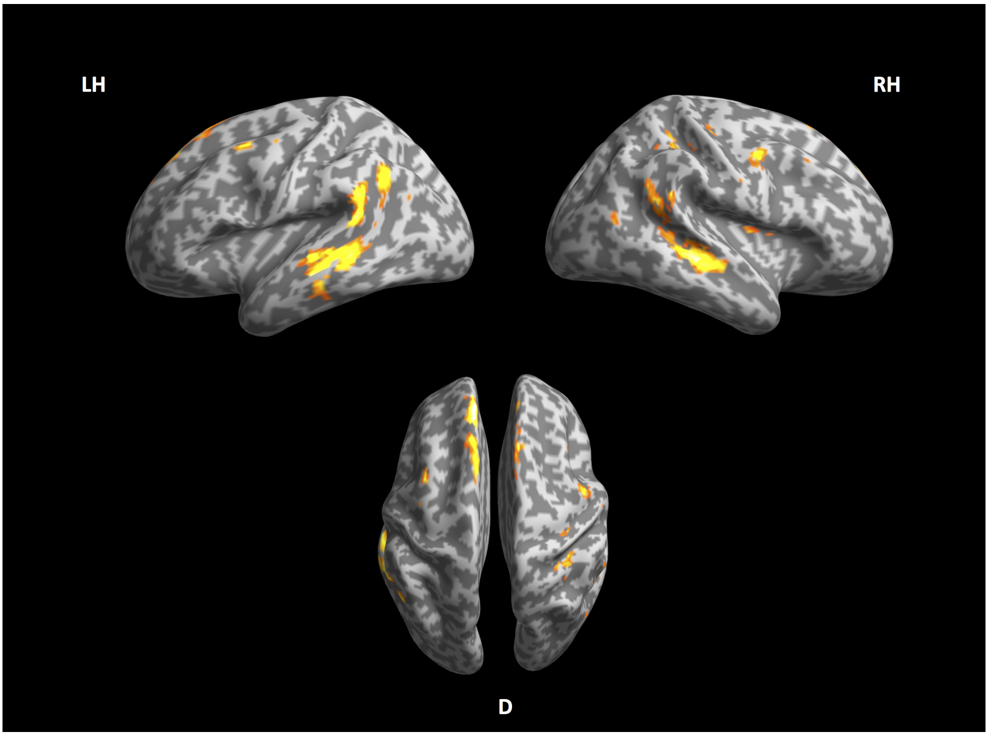


Figure 3.3: Syntactic stream. Left, right and ventral view of inflated cortex plot of the syntactic stream.

Region	**MNI**	**size**	**t-value max**
left middle superior frontal gyrus	-6 34 56	1549	6.50
left precentral sulcus	-42 6 54	267	6.19
left middle temporal gyrus & sulcus	-64 -50 14	1715	5.64
left and right cerebellum 9	-20 -46 -36	662	5.07
right middle temporal sulcus	48 -32 -2	983	5.51
right angular gyrus	64 -50 24		5.49
right superior temporal sulcus	54 -22 -6		4.71
right putamen	24 -2 6	649	6.04
right amygdala	22 2 -8		4.46
right precentral sulcus	52 -2 46	119	4.31

Table 3.7: Significant effects of syntactic stream perplexity.

3.5.3 Phonological stream

Fig 3.4 and Table 3.8 refer to the network of the phonological stream. This stream involves the right Heschl's gyrus (r-HG), and right superior frontal gyrus (r-SFG) together with the supplementary motor area (r-SMA). Other areas activated to this contrast are the left insula, the left angular gyrus (l-AG), the left inferior parietal lobule (l-IPL), and bilateral portions of the middle temporal gyrus (rl-MTG). The phonological stream was the only level where perplexity and surprisal did not give comparable results.

Region	**MNI**	**size**	**t-value max**
left insula	-36 8 -18	123	4.75
left angular gyrus	-40 -56 40	1507	4.78
left inferior parietal lobule	-42 -44 44		4.35
left posterior mid temporal gyrus	-42 -64 16	137	3.71
right Heschl's gyrus	50 -12 4	443	4.63
right Heschl's gyrus	40 -24 12		4.46
right posterior mid temporal gyrus	42 -64 16	950	4.14
right angular gyrus	56 -56 24		4.09
right superior frontal gyrus - SMA	20 16 62	202	4.49

Table 3.8: Significant effects of phonological stream perplexity.

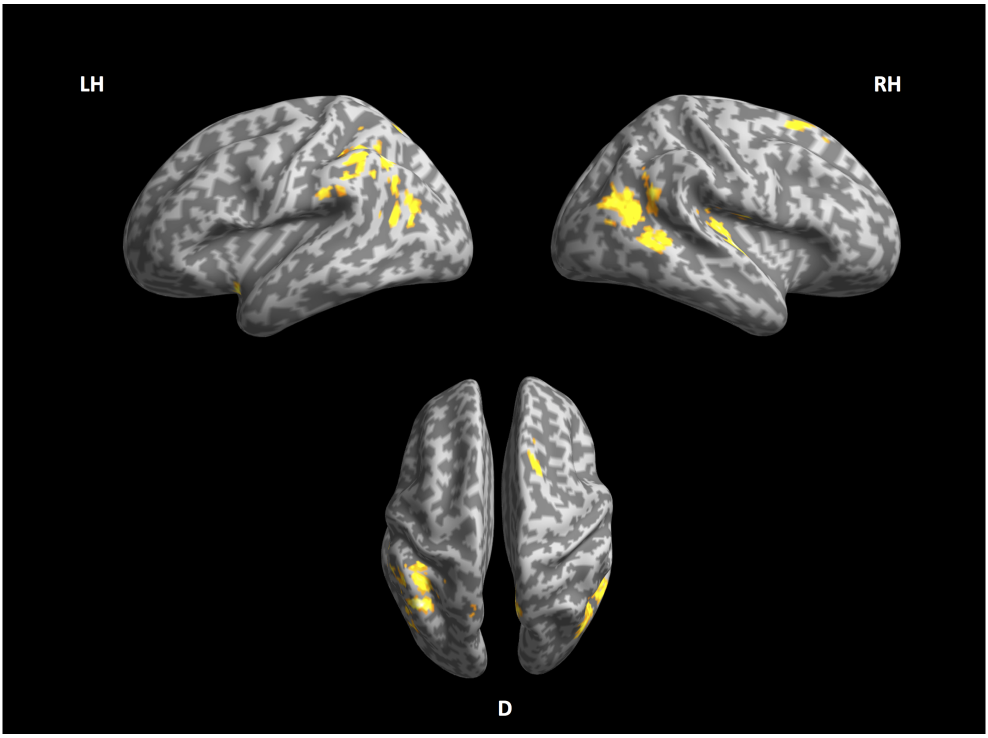


Figure 3.4: Phonological stream. Left, right and ventral view of inflated cortex plot of the phonological stream.

3.5.4 Overlap

The results obtained from lexical, syntactic, and phonological perplexity allowed me to analyze possible overlap at the cortical level for all three regressor streams. Table 3.9 contains the name of regions whose activity is significantly explained by more than one regressor. I computed the overlap between the activation maps relative to the lexical and the syntactic stream, the lexical and phonological streams, and the syntactic and phonological stream by taking only those voxels that are significantly activated for both regressors as described in the sections above. This is equivalent to performing a conjunction analysis, more specifically a test of the 'conjunction null', effectively looking for statistical significance in both contrast maps as testing a logical AND (Nichols et al., 2005). I also looked for voxels shared by all three streams.

It is interesting to note that although some degree of anatomical overlap exists among all possible pairs of regressors, there is no area that is significantly activated for all three streams together. What is also worth noting is that the lexical and syntactic regressors are

Streams	**regions**	**$\sim$MNI coordinates**
Syntax ∩ Lexical	right STS	$54\ -25\ 0$
	left posterior MTG	$-58\ -51\ 11$
Syntax ∩ Phonology	right AG	$59\ -56\ 29$
	left AG	$-53\ -64\ 31$
Lexical ∩ Phonology	right middle STG	$44\ -29\ 13$
Syntax ∩ Lexical ∩ Phonology	∅	∅

Table 3.9: Areas of overlap between the streams regressors.

both processed in the posterior portions of the bilateral middle temporal gyrus, bordering the posterior superior temporal gyrus. Moreover, lexical information and phonology seem to share activity in the central banks of the superior temporal gyrus, but not directly in Heschl's gyrus, which confirms its selectivity for the phonological stream only. The overlap regions are shown in cyan (lexical and syntactical streams) and in violet (syntactic and phonological streams) in Figures 3.5 and 3.6.

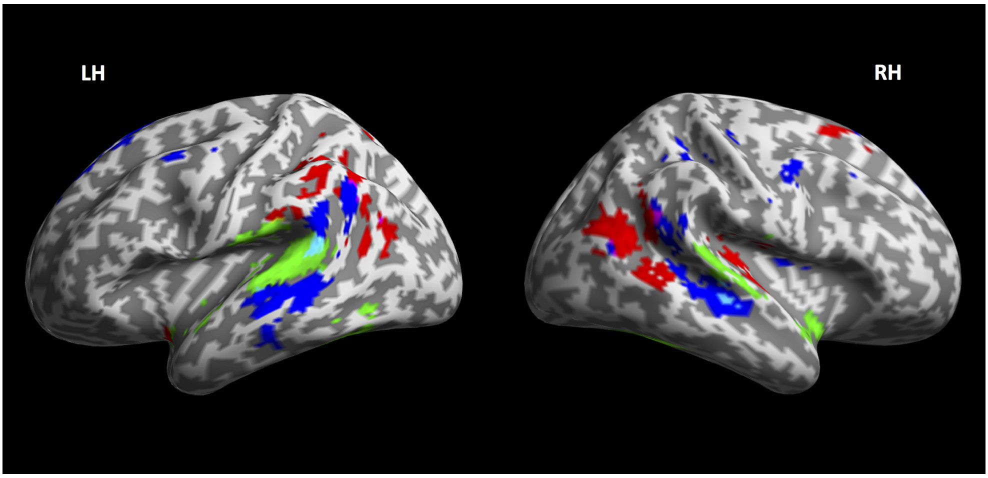


Figure 3.5: Streams comparison and overlap. Inflated cortex view of the maps of the lexical (green), syntactic (blue), and phonological (red) streams. In this view, the overlap between lexical and syntactic streams is particularly evident in the right Middle Temporal Lobe and in the left posterior Superior Temporal Gyrus (cyan). Overlap between syntactic and phonological streams is also evident in the bilateral Angular Gyrus (violet).

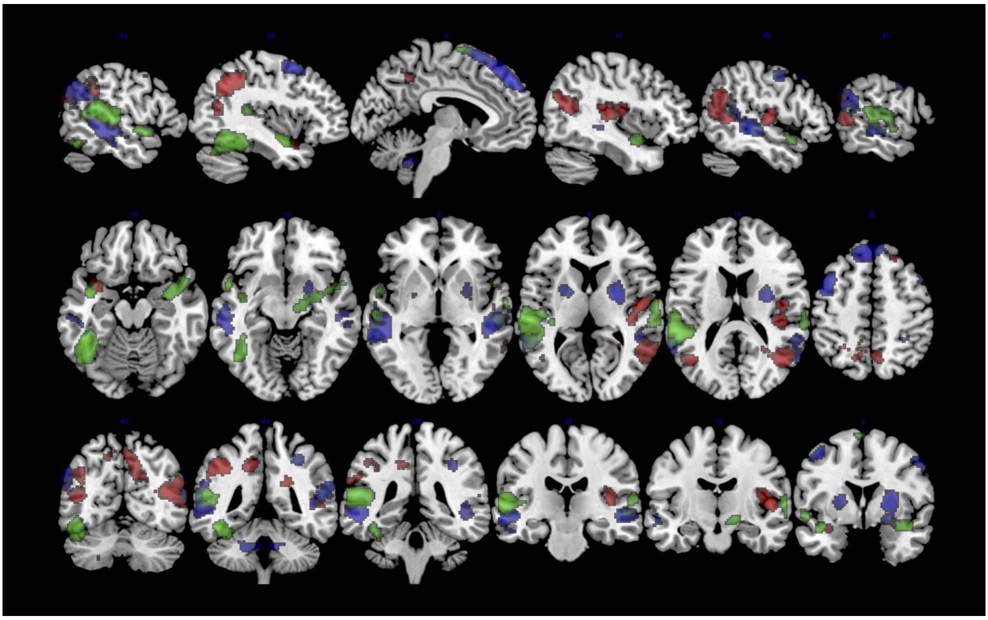


Figure 3.6: Streams comparison and overlap. Sagittal, axial, and coronal view of the maps of the lexical (green), syntactic (blue), and phonological (red) streams.

3.6 Discussion

The results reported above outline a set of cortical networks that are separately activated for each of the three types of information under investigation – lexical, syntactic, and phonological -– confirming the hypothesis that language processing can indeed be decomposed into different streams corresponding to different subdivisions of the language network. No area shows selectivity for all three streams, and only limited sets of voxels show overlap between pairs of streams.

3.6.1 Division of labor in the temporal cortex

The temporal lobe shows a distribution between the three streams that sees the lexical information primarily concerning the infero-lateral regions, syntactic information the mid-lateral regions, and phonological information finding its hub in the middle superior temporal plane. Areas posterior to the perisylvian cortex, between AG, SMG, and IPL, display a similar gradient, with more rostral voxels selective to lexical information and more caudal ones phonological information, with selectivity to PoS-related information in the middle.

3.6.2 Phonological stream

The phonological stream seems to involve activity in the temporal cortex only in regions close to the transverse gyrus (Heschl's gyrus), especially in the right hemisphere. This area is the central hub of auditory processing (Mendoza, 2011). Although the phonological regressor is built on a level of annotation that is close to the actual perceptual structure of the words, it is not directly built on the auditory properties of the stimulus, making these observed results both surprising and interesting. In addition to this, the phonological stream activates the supplementary motor area (SMA), which has been suggested to be involved not only in speech production (Alario et al., 2006) but also in speech processing (Willems & Hagoort, 2007; Hertrich et al., 2016). Phoneme perplexity did not return activation in the premotor cortex, an area that has been associated with speech production and perception. Nonetheless, activation of the premotor cortex in response to phonological load is not a general finding in the literature. Tremblay & Small, 2011; Tremblay et al., 2016 have suggested that premotor cortex activation during speech processing may only be observed under tasks presenting particularly difficult conditions. In line with this position, Sato et al., 2009 have shown that stimulating the premotor cortex only has an effect on a complex speech perception task. Similarly, premotor activity has been found to be modulated by syllable complexity during speech production but not during speech perception (Tremblay & Small, 2011).

3.6.3 Lexical and syntactic streams

Middle temporal gyrus

Dronkers et al., 2004 have suggested that the posterior MTG plays a role in retrieving lexical and syntactic properties of incoming words from long-term memory. Hagoort, 2013a suggests that MTG might be important for the retrieval of the syntactic frames (as well as other lexical information) from the mental lexicon, which are then combined in the left inferior frontal cortex (Indefrey & Cutler, 2004). This intuition is corroborated by the present results, which show that activity in this area, although mainly explained by the syntactic regressors employed in these analyses, displays an overlap between the syntactic regressor and the lexical stream. The work of Snijders et al., 2008 showed that Dutch noun-verb homonyms (grammatical category ambiguity) increased activity in the posterior MTG. This study also reports that grammatically ambiguous sentences activated not only the posterior MTG but also the precentral gyrus, an area that is also observed in the present analyses.

Although PoS perplexity intends to model syntactic processing and appears to be a reasonable correlate of the syntactic stream, my analysis shows only a marginal correlation between this regressor and the activity observed in the anterior temporal cortex. Studies from

Pallier et al., 2011, Obleser et al., 2011 and Brennan et al., 2012 suggest that this region is sensitive to the syntactic complexity of the input sentence. Brennan and colleagues used a similar experimental paradigm to the one adopted in this chapter. They had participants listen to a segment of a novel (Lewis Carroll's *Alice in Wonderland*) and looked for areas of which the activity correlates with the number of so-called syntactic building operations at each time point, representing the number of non-terminal phrases that are completed by the presentation of each word. This measure is based on a hierarchical treatment of syntax, whereas I intended to model syntactic processing in a purely sequential manner. These two results can be reconciled by considering syntactic processing as underpinned by both sequential probabilistic machinery (captured by PoS-perplexity) and hierarchical structure building.

Inferior temporal cortex

Activity in the inferior and lateral portions of the left temporal cortex is better explained by the lexical regressor, which is likely to be a central hub of the lexical stream. Willems et al., 2016 observed the same result using the same dataset, nonetheless interpreting it as activity in the visual word form area (VWFA, Cohen, 2002). Their explanation is that word prediction can account for the pre-activation of the upcoming word form in the sequential sentence processing. This explanation is not the only possible one. For instance, Price & Devlin, 2003 points out that the cortical region corresponding to VWFA is active in normal subjects also during tasks that do not engage in visual word form processing. On the other hand, if activity in the left Inferior Temporal cortex and specifically in VWFA truly reflects word form prediction, I would have also expected phonological perplexity to show selectivity in this region. Phonological perplexity, computed on the phonemic structure of every single word, seems intuitively a closer proxy for the form of a word. Although computed on the phonemic transcription of the words, the relation between phoneme and grapheme in Dutch is at least somewhat regular, making visual and phonemic structure intuitively close. Nonetheless, as explained above, this model does not predict activity in ITC better than word-based perplexity. This suggests indeed that the coupling between the later regressor and activity in this region reflects lexico-semantic rather than word form information.

That the lexical regressor, computed in terms of trigram statistics in the co-occurrences of words, is a correlate of lexical semantic processing is strengthened by the outcome of a meta-analysis of 120 functional neuroimaging studies (Binder et al., 2009; Binder & Desai, 2011). The meta-analysis showed that the lateral and ventral temporal cortex is among the main nodes of the semantic processing network. This interpretation is supported by studies that reported consistent correlation between lexical semantic models and brain activity in ventro-temporal cortex (Anderson et al., 2013, 2015).

3.6.4 Left inferior frontal gyrus

None of the perplexity-based regressors returned significant activation in the left inferior frontal gyrus (l-IFG). While l-IFG is an important node in the neural language network, its involvement and potential role during language comprehension has been the subject of considerable debate.

One line of work starting with Thompson-Schill et al., 2009 has argued that the role of this area is better characterized as a general, not language-specific one, and involved in 'selection' or – more generally – 'cognitive control'. Another approach has stressed the role of the area in structural processing, both in a hierarchical and sequential fashion (Grodzinsky & Santi, 2008; Bornkessel-Schlesewsky & Schlesewsky, 2012). Nonetheless, not all results seem to support this view. Brennan et al., 2012 and Brennan et al., 2016, for instance, found that syntactic complexity did not correlate with l-IFG activity, which seems at odds with some previous findings that did observe l-IFG activation in response to syntactically hard to parse sentences. The fact that l-IFG was not detected in Brennan's work and in the work presented in the present chapter might be due to methodological reasons. Both Brennan and I used naturalistic stimuli and correlation between brain imaging data and stimuli properties (stochastic in my case, hierarchically structural in the case of Brennan and colleagues). The literature advocating the role of l-IFG in processing is dominated instead by paradigms comparing carefully constructed sentences, for instance, syntactically ambiguous vs. unambiguous (Snijders et al., 2008), or grammatical vs. ungrammatical (Friederici et al., 2003; Herrmann et al., 2012).

In the scope of the present chapter, I cannot draw any strong conclusion regarding l-IFG on the basis of its 'non-activation'.

3.7 Conclusions

The analyses presented in this chapter have shown that the stochastic sequential processing paradigm is indeed a powerful formalism able to predict neurobiological correlates in areas belonging to the language processing network, also when applied to sub-lexical (phonemic) and syntactic (part of speech) levels. Previous work has demonstrated that language processing can be characterized as a stochastic process computed on sequences of words and that measures of stochastic perplexity are good predictors of brain activity in language-sensitive cortical areas.

Word-based (lexical), part of speech-based, and phoneme-based perplexity distinctively predict activity in largely separated cortical networks in the temporal, inferior parietal, and perisylvian cortex of subjects listening to naturalistic linguistic input.

These results appear to confirm the intuition that language is processed in parallel by distinct networks sensitive to different sources of information, including at least the ones tested here: phonological, lexical, and syntactic.

4

Dependency and Phrase-Structure Processing in the Brain

The language system in the brain is not involved merely in the processing of the sequential properties of the stimulus. Other structural analyses are required in order to interpret language. Finding the structure of a sentence — the way its words hold together — is a fundamental step in language comprehension. Several brain regions, including the left inferior frontal gyrus, the left posterior superior temporal gyrus, and the left anterior temporal pole, are supposed to support this operation. The exact role of these areas is nonetheless still debated. In this paper, I investigate the hypothesis that different brain regions could be sensitive to different kinds of syntactic computations. I compare the fit of phrase-structure and dependency-structure descriptors to activity in brain areas using fMRI. The results show a division between areas with regard to the type of structure computed, with the left ATP and left IFG favoring dependency structures and left pSTG favoring phrase structures.

This chapter is based on:

Lopopolo, Alessandro, Antal van den Bosch, Karl-Magnus Petersson and Roel M. Willems. 2020. Distinguishing syntactic operations in the brain: Dependency and phrase-structure parsing. [under review].

The previous chapter began with the observation that a sentence is first and foremost, a sequence of symbols. I also hypothesized, and then demonstrated, that sequential properties of the stimulus drive part of the neural activity involved in language processing and that they can be decomposed in several distinct levels of information treating the stimulus as a sequence of phonemes (the stimulus was spoken), word forms and grammatical classes.

In this section, I move away from the sequential properties of the stimulus, and its putative levels of analysis, and instead I focus on the type of structure the brain is supposed to construct on top of the received stimulus. As I will argue in the first sections of this chapter, understanding a linguistic signal – let us say, a sentence – cannot rely solely on a list of words or grammatical categories, but requires the discovery of relations between words, even distant in space, describing the argument structure of the sentence. The argument is that the brain computes a hierarchical structure – syntactic if you will – that groups words together according to some structural criterion.

The fact that there are pieces of evidence for both sequential and hierarchical processing in the brain is a known fact (Brennan et al., 2016; Nelson et al., 2017). In this chapter, I address a couple of related questions instead. By adopting two distinct forms of syntactic structures, one that directly captures thematic relations between the components of the sentence and another that emphasizes *phrasal structure* instead, I investigate whether syntactic analysis in the brain can be in turn decomposed in sub-networks sensitive to different types of structural relations.

4.1 Brain areas underpinning syntactic analysis

Sentence processing involves at least two operations: the retrieval of the meaning of single linguistic units from semantic memory and the computation of the meaning of the structures derived from the combination of these – more basic – units. This second sub-process likely requires the contribution of some sort of structural analysis, i.e., the analysis of the syntactic configuration of the words making up the sentence. In this section, I will review and motivate a selection of cortical areas that – not uncontroversially – seem to support structural analysis.

The literature reports the involvement of a network of mostly left-lateralized cortical regions, including the left inferior frontal gyrus (IFG), the left posterior superior temporal gyrus (pSTG), and the left anterior temporal pole (ATP). There is, however, controversy concerning which brain areas are crucially involved in syntactic processing. The effective involvement and the division of labor of these cortical areas, is still a matter of ample debate. Not all studies on sentential processing arrive at the conclusion that IFG and pSTG are in fact involved in syntactic processing. Although there is a larger consensus regarding the involvement of the left ATP in sentence processing, the left ATP has been equally linked to two types of sub-processes: syntactic composition – i.e., the ability to derive larger units by combining smaller ones – and lexical semantic memory representation.

A considerable body of literature does report left IFG and left pSTG activation during syntactic processing as opposed to a baseline, usually consisting of random sequences of words (Caramazza & Zurif, 1976; Friederici et al., 2005; Tyler et al., 2008; Pallier et al., 2011; Zaccarella & Friederici, 2015; Zaccarella et al., 2015). These findings are contradicted by a set of studies that do not report activity in left IFG and left pSTG (Humphries et al., 2006; Rogalsky & Hickok, 2008; Bemis & Pylkkänen, 2011), although using similar paradigms to the above-mentioned studies. Moreover, doubts about the effective involvement of these areas in syntactic processing are cast by neuropsychological observations. For instance, lesions to the IFG lead to what is clinically known as Broca's aphasia. These aphasic patients do not perform significantly different from healthy controls on grammaticality judgment (Linebarger et al., 1983; Wulfeck & Bates, 1991). Similarly, lesion analyses seem to point towards a lack of effect of lesions located in IFG and pSTG to the performance in basic sentence comprehension (Dronkers et al., 2004; Thothathiri et al., 2012). These studies do not target specific syntactic structures or syntactic structure processing directly. Nonetheless, both tasks – grammaticality judgment and sentence comprehension – are likely to require the computation and the analysis of the syntactic structure of the presented stimuli.

Acknowledging this inconsistency in the literature regarding the involvement of frontal and posterior temporal regions, Matchin et al., 2017 proposed the hypothesis that the left IFG and pSTG may not play a necessary role in syntactic processing. Instead, they claim that these areas are involved only in top-down syntactic prediction supporting compositional syntactic operations in left ATP and left AG.

Previously, Pallier et al., 2011 analyzed the activity recorded during natural language sentence and jabberwocky sentence reading. Searching for brain regions where activation positively correlated with the size of linguistic constituents, their results isolated a network of left-hemispheric regions that could be dissociated into two major subsets. Inferior frontal and posterior temporal regions showed constituent size effects regardless of whether actual content words were present or were replaced by pseudowords (jabberwocky stimuli). On the other hand, regions in the temporal pole, anterior superior temporal sulcus, and temporo-parietal junction showed constituent size effect only in the presence of lexico-semantic information, suggesting that they may encode sentence-level semantic compositionality.

The left ATP plays a central role in linguistic processing and has been reported as a central hub for lexical, semantic, and syntactic compositionality. A large literature has pointed to the involvement of the left ATP in the processing of sentence and phrasal structure. By contrasting the activity recorded during the reading of sentences and of word lists, works such as Mazoyer et al., 1993, Stowe et al., 1998, Friederici et al., 2000, Humphries et al., 2006, and Humphries et al., 2007 reported an increase in activity in ATP for sentence comprehension as compared to word lists. The role of ATP in processing composition is confirmed by another series of studies that focused on more specific types of syntactic structures. Rather than looking at sentences as a whole, these analyses focused on simple phrasal processing, consisting of the composition of adjectives and nouns (e.g. *red apple*) (Baron et al., 2010; Baron & Osherson, 2011; Bemis & Pylkkänen, 2011, 2013). These results are also confirmed for a wider range of phrasal and syntactic compositional types and cross-language by Westerlund et al., 2015, and across visual and auditory modality by Bemis & Pylkkänen, 2013.

Next to its involvement in syntactic processing, the left ATP is also considered central in semantic memory, a putative subcomponent of long-term memory storing information about the meaning of linguistic units. The first and most compelling proof of this role of ATP is given by the studies on semantic dementia (SD), in which patients showing atrophy of ATP show a significant impairment in their ability to retrieve and recognize concepts (Hodges et al., 1992, 1995; Mummery et al., 2000; Rogers et al., 2004). This is confirmed also by a large neuroimaging literature (Gauthier et al., 1997; Tyler et al., 2004; Bright et al., 2005; Moss et al., 2004; Rogers et al., 2006). These findings were summarised by Patterson et al., 2007 and led to the formulation of the hub and spokes model, which posits that concepts are represented by a network of sensorimotor representations converging in the left ATP which acts as a hub collecting and controlling modality-specific features in order to produce supra-modal representations.

Following the studies on sentential and phrasal processing, and Patterson's hub and spoke model of semantic memory, it appears that ATP could play a role in two distinct kinds of composition: one merging word into larger structures (phrases and sentences), and one composing words out of more basic semantic features, possibly grounded in sensory-motor representations. This led Westerlund & Pylkkänen, 2014 to compare the involvement of ATP between tasks requiring syntactic and lexical semantic processing, concluding that the two processes might indeed be substantiated by the same cortical mechanism.

4.2 What form of syntax?

In the previous section, we saw how the debate on the cortical involvement during structural sentence analysis generally points to areas in the left inferior frontal (IFG), left superior temporal and left anterotemporal (pSTG and ATP) regions. The question I address in this study is whether they are differently involved in specific syntactic computations.

Recent studies such as Brennan et al., 2016 and Nelson et al., 2017 report phrase-structure correlating with activity in the left ATP. Since other studies showed instead sensitivity to more basic structures such as adjective-noun, adverb-verb, adverb-adjective, and argument-predicate (Westerlund et al., 2015), I question whether phase-structure is the only formalism that can characterize structural processing, especially in the left ATP.

In this chapter, I compare **phrase-structure grammars** (**PSG**) (Chomsky, 1957, 1965; Borsley, 1998) and **dependency grammars** (**DG**) (Tesnière et al., 2015; Mel'čuk, 1988; Nivre & Kübler, 2009) as two different hypotheses regarding the type of structure the brain computes as part of sentence comprehension. The two grammars differ in a number of aspects. DG builds structures solely on the words and on binary relations holding between them, whereas PSG rely on grouping words in phrases that can, in turn, be grouped in larger phrases introducing a hierarchical structure composed by both surface forms (the words of the sentence) and non-observable abstract nodes that are assumed to be computed by the human brain.

In order to test the different contribution of phrase and dependency structure, and ultimately assess the role of syntactic analysis in language processing in the brain, I introduce syntactic descriptors that can be derived from these two types of parses. These descriptors relate to the structure-building processes that allegedly take place if the brain is involved in syntactic processing. These consist in the number of non-terminal (non-surface) nodes attached to each word in a sentence in case of phrase-structure grammar, and the number of terminal (surface) nodes attached to it as dependents in case of the dependency grammar. It is important to keep in mind, that my aim is not to prove that one grammar is a better formalism than the other. I intend to investigate whether and how the language network in the brain is sensitive to measures derived from them.

Inspired by the previous literature, I conducted a region of interest (ROI) analysis focusing on the left IFG (pars opercularis, triangularis, and orbitalis separately), the left ATP, and left STG. I fitted separate linear mixed-effect (lme) models predicting the activity recorded in these areas during naturalistic language listening, using as regressors of interest the structural measures mentioned above. Note that the regressors specify the amount of syntactic processing at each word in the stimuli. These analyses allowed me to identify which area is more sensitive to which type of structural description (PSG or DG). I then conducted a psychophysiological interaction (PPI) analysis investigating how the interaction between each of the ROI's and the rest of the brain is modulated by its preferred structural description as from the previous lme-analysis.

In a nutshell, the results indicate that PSG significantly predicts activity in the left STG, but not, as reported in other studies (Brennan et al., 2016; Nelson et al., 2017), in the left IFG and in the ATP. Activity in these areas is better explained by DG predictors. Moreover, I observed how the activity in left STG, modulated by PSG measures, might drive the activity in the left ATP. These observations suggest a division of labor in the cortical areas processing syntax among areas involved in phrase structure analysis and others possibly involved in dependency analysis. More importantly, the results show that the left ATP is sensitive to DG measures, rather than PSG measures.

4.3 Syntactic structures

The assumption introduced earlier is that in order to interpret a sentence, the human brain has to to establish relations between the words that compose it. For instance, words alone, in isolation, cannot convey the full description of a situation or a state. The following list of words – *chapter*, *you*, *this*, and *read* – becomes a suitable description of the action you are performing now only if the relations that the predicate *read* entertains with the subject *you* and object *this chapter* – in turn substantiated by the relation between determiner *this* and noun *chapter* – are established by the brain. In other words, the interpretation of a sentence does not only depend on the meaning and the grammatical category of the words that compose it, but also on their syntactic relations. The set of relations that hold between the words constitutes the syntactic structure of the sentence. The way we can describe these relations and structures can formalized in terms of grammars. In the remainder of this section I will describe the two types of grammar studies in this chapter: phrase structure (PSG) and dependency grammar (DG).

Given a sentence, both PSG and DG produce a hierarchical structure linking or grouping the words in a structure rooted in a governing node (the root node). The main difference between the two is that PSG assumes the existence of phrase structures grouping and governing pairs (if the parse tree is binary) of words, whereas DG relies only on word pairings linked by syntactic relations. At a high level of abstraction, what sets a DG structure apart from one derived according to PSG is the fact that DG structures are flatter than PSG structures because they lack phrasal constituents. The structure only consists of the words in a sentence and an associated set of directed binary grammatical relations that hold among them. The only nodes in the DG structure are terminal nodes corresponding to surface lexical items as they are encountered by the human reader; no non-terminal, non-observed abstract nodes are introduced.

In the following section, I expand on the fundamental differences between PSG and DG both from a structural and theoretical point of view.

4.3.1 Phrase structure

The PSG parse of a sentence is a tree structure composed by nodes: terminal and non-terminal. Non-terminal nodes correspond, usually, to phrases and therefore are assigned labels corresponding to syntactic phrasal categories such as Noun phrase (NP), Verb phrase (VP), Adverbial phrase (AP), and Determiner phrase (DP). Terminal nodes, instead, are the *leafs* of the tree and correspond to the surface forms of the parsed sentence, i.e., its words.

If the tree is binary – in the definition of phrase-structure parse, I adopt only binarized trees – phrasal nodes can have a maximum of two child nodes that can be either other phrasal nodes or leaf nodes (words). A parent node can only consist of a phrasal node; it is also referred to as non-terminal. Words can only be children of non-terminal phrasal nodes and are referred to as terminal or leaf nodes because they are not hierarchically higher than any other node. Besides phrasal and leaf nodes, the phrase-structure parse also contains a root node. A root node is a node that is not a child of any other node.

There is only one one root node in the parse, and it corresponds to the category S, governing the sentence as a whole.

(4) The man saw a brown dog in the park.

As an example, as displayed in Figure 4.1, the parse of sentence 4 contains eight labeled phrase structures, including S, and constitutes a nested binary-branching tree. The words of the sentence (*the*, *man*, *saw*, *a*, *brown*, *dog*, *in*, *the*, and *park)* correspond to the terminal nodes. Following the structure of the parse tree in a top-down fashion: S branches into a NP and VP (noun and verb phrase, respectively). The left-hand child (NP) is composed of a determiner leaf node *the* and a noun *man*; whereas the right-hand child of S (VP) has in turn as left-hand child a terminal node (the finite verb *saw*) and as its right-hand side child another noun phrase (NP). This last NP branches of in a another NP and in a prepositional phrase (PP). These two last phrases both split in a left-hand terminal child (respectively *a* and *in*) and in a NP as right-hand child. The latter two are both composed of terminal nodes (*brown*, *dog*, *the*, and *park*).

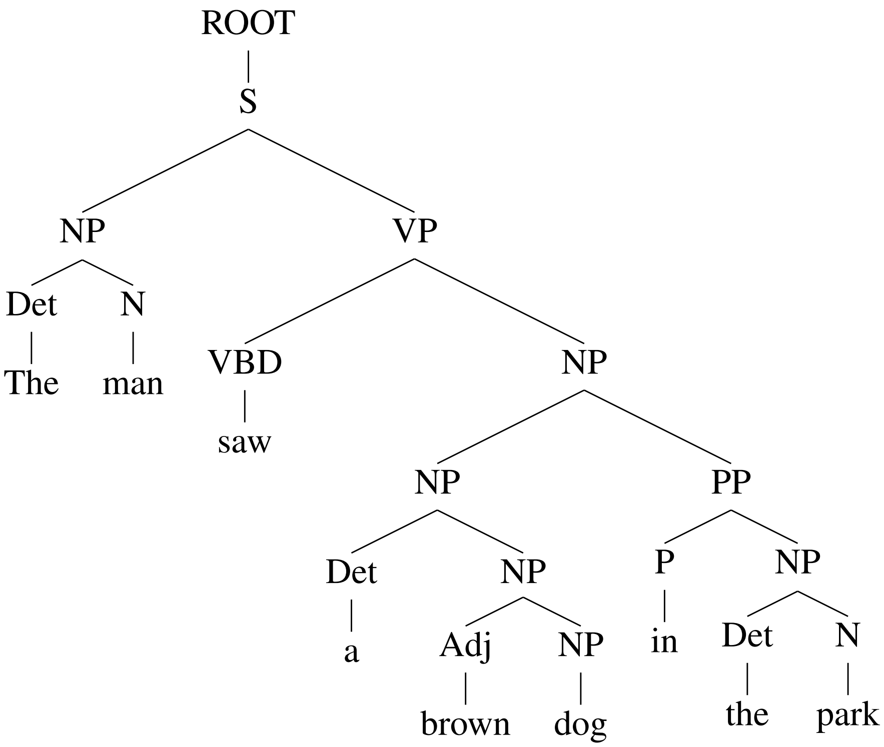


Figure 4.1: Phrase-structure parse of Sentence 4, taken from Chapter 1.2.2.

4.3.2 Dependency structure

When computing DG parses, sentences are described solely by binary relations between the pairs of words composing them. These relations, or dependencies, correspond to grammatical functions (Tesnière et al., 2015; Mel'čuk, 1988; Nivre & Kübler, 2009). A DG parse is also anchored to a root node that directly governs the main finite verb of the sentence, which, in turn, can be considered the structural hub of the sentence itself.

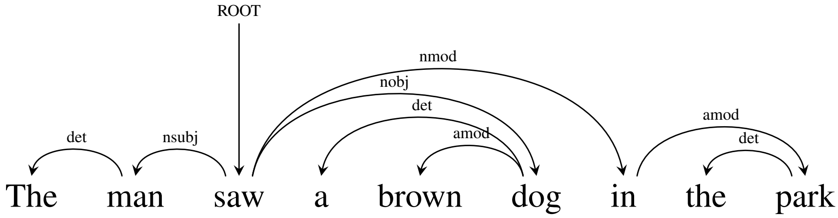


Figure 4.2: Dependency parse of Sentence 7, taken from Chapter 1.2.2.

Figure 4.2 display the DG parse of Sentence 7 in terms of typified head-dependent relations: the main verb (*saw*) acts as head for *man* and *dog*, with which the verb is in a **subject** and **object** relation respectively. A dependent of one dependency relation can in turn be the head of another. For instance *dog* is both dependent of *saw* and head of both *brown* (via a **modifier** relation) and *a* (via a **determiner** relation). As it is evident for these examples, dependencies can be instantiated between words far apart in the sequential structure of the sentence.

Relations that hold between words are captured in structurally different manners by dependency structure, and phrase structure parses of the same sentence. Take for instance the relation between *saw* and *dog*, respectively the main verb and the direct object in Sentence 4. As it is apparent by simply looking at the graph path between these two items in Figure 4.3.a and 4.3.b, the dependency structure directly captures their predicate-object relation by mean of a simple directed edge, whereas PS relies on three intervening noun phrases and a governing verb phrase.

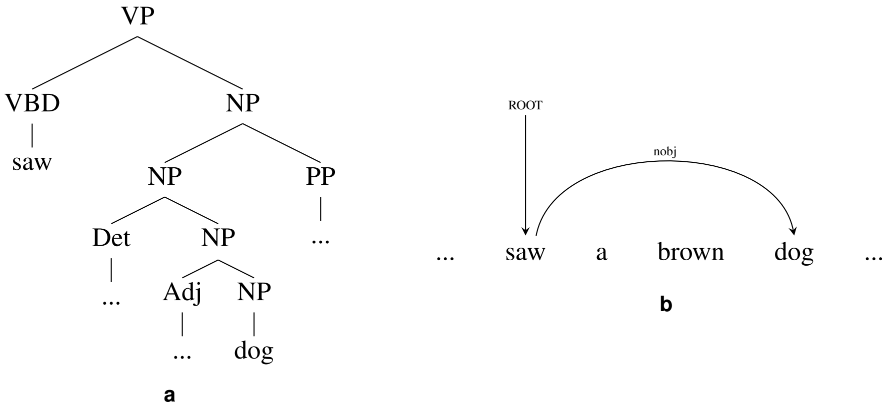


Figure 4.3: Comparison between the structures mediating the relation between the sentence's main verb (*saw*) and its object (*dog*) in the phrase-structure (a) and dependency structure (b) parses of Sentence 7.

4.4 Materials and methods

4.4.1 fMRI data

For the purpose of this study, I have used the fMRI narrative brain dataset described in Chapter 2 Section 2.1. The dataset consists of data of 24 native Dutch speakers collected while the participants were asked to listen to the spoken presentation of three Dutch narrative texts.

4.4.2 Syntactic measures

Both dependency and phrase-structure parses of the sentences composing the stimulus texts are derived using a computational parser developed and trained for Dutch (ALPINO, Noord, 2006). Alpino is considered the state of the art standard parser for Dutch, and it is consistently being used for natural language processing applications requiring syntactic analyses of Dutch texts. The output of Alpino is able to return sentence parses consistent with the principle of phrase-structure and dependency grammars as delineated in Section 4.3. Moreover it is able to generate both these two types of parses within the same framework, making it convenient and allowing us to avoid inconsistencies derived from using different parsers built and trained on different data. From these parse structures, I derive measures approximating the operations performed in order to integrate each word in the syntactic structure computed at the point of its presentation. The next sections will describe these measures in detail.

Dependency parse

In order to describe the dependency structure of a sentence, the ALPINO parser creates a structure composed by dependency triplets consisting of: a head word, the type of dependency relation, and its dependent word. A parse is produced for each sentence independently; therefore, no relation can be assigned between words belonging to different sentences.

In order to describe the operation required to integrate a word at a time in the incrementally built dependency structure of the sentence, I adopted the number of left-hand side relations entertained by each word. As described in Section 4.3.2, every word in a sentence entertains at least one relation with another word in the same sentence. Every non-final and non-initial word can have relations with a variable number of other words on its right and its left. Logically, a sentence-initial word can only have relations with words to its right, and a sentence-final word can only be linked to words on its left. In order to quantify the operations required to integrate a word w in the structure constructed up to its presentation, only relations with a head and possible dependents on the left hand-side of w are counted. In other words, from the Dependency Structure of a sentence, I count the number of left hand-side edges for each word w in the sentence (dependency structure left relations or DSlrels, see Table 4.1).

	The	man	saw	a	brown	dog	in	the	park
DSlrels	0	1	1	0	0	3	1	0	1

Table 4.1: Number of left-hand dependency relations (DSlrels) per word w in the example Sentence 4.

For example, the word *dog* in the sentence has two dependent relations with two words to its left (*a* and *brown*), no dependents to its right, but one head to its left (*saw*). The word *park*, being sentence-final, does not have any links on its right, but it has one head (*in*) and one dependent (*the*) to its left.

Phrase-structure parse

As for the dependency parse, the texts of the three stories presented to the participants were fed to ALPINO toolbox for Dutch natural language processing in order to generate this time a phrase-structure parse for each sentence (Noord, 2006).

In order to quantify the number of syntactic operations per word required to construct a phrase-structure parse of the input sentence, I measure the number of closed phrase structures allowed after the introduction of each novel word (PSxps). Such a measure is computed by considering whether a word or phrase is a right-hand or left-hand side child of its parent phrasal node. In case the word in question is the right-hand side child, the parent phrasal node is considered complete and therefore closed. This proceeds recursively, evaluating whether a closed phrasal node is in its turn the right-hand side child of a higher-order parent phrasal node, allowing it to be closed. For instance, according to the phrase-structure parse of Sentence 4 (Figure 4.1), the first instance of the word *The* is the left-hand-side child of an NP structure; for this reason, this NP is not complete and cannot be closed. Therefore, the value of PSxps for *The* is 0. On the other hand, *man* is the right-hand-side child of the same NP, and therefore this phrase structure can be closed at this word position, allowing the assignment of value 1 to *man*. Following the same reasoning, *dog* is the left-hand side child of another NP, allowing for its closure. This last NP is, in turn, the left-hand side of a higher NP structure. Therefore the word *dog* is assigned value 2 because its presentation allows for the completion of two nested phrase structures. Table 4.2 contains the PSxps values for the whole Sentence 4.

This measure is computed under the following simplifying assumptions: that phrase-structure trees are binary (i.e., as explained above, that they can have only 2 children), and that parsing proceeds incrementally left-to-right.

	The	man	saw	a	brown	dog	in	the	park
PSxps	0	1	0	0	0	2	0	0	5

Table 4.2: Number of closed phrase structures (PSxps) per each word w in the example Sentence 4.

4.4.3 Controlling for lexical frequency and word surprisal

To control for other factors known to influence brain activation during language comprehension, I added log-transformed lexical frequency and surprisal as covariates to the analysis (Willems et al., 2016; Lopopolo et al., 2017).

Log2-transformed lexical frequency per word was computed using the Subtlex NL corpus (Keuleers et al., 2010). Surprisal was computed from a second-order Markov model, also known as trigram model, trained on a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the Dutch Corpus of Web (Schäfer & Bildhauer, 2012). Surprisal of word w_t is the negative logarithm of the conditional probability of encountering w_t after having read sequence w_{t-2}, w_{t-1}, or: $-\log P(w_t|w_{t-2}, ...w_{t-1})$. The computation was performed by the SRILM toolbox (Stolcke, 2002).

4.5 Analyses

The main analysis consists of a region of interest-wise linear model fitting using as predictors the syntactic structure measures described in the previous sections, together with lexical frequency and surprisal as regressors of no interest.

Besides the ROI-analyses, I also conducted a whole-brain psychophysiological interaction (PPI) analysis. The former was performed in order to test the interaction between the regions of the language network and the rest of the brain with regard to the type of syntactic structure considered in this study. The later was instead performed in order to have a wider – less biased – view of the possible division of labor between dependency and phrase-structure structure parsing in the brain.

4.5.1 ROI analysis

I chose six separate left-hemisphere anatomical regions of interest (ROIs) to selectively test the contribution of the two syntactic measures as predictors of BOLD activity. These regions were: superior temporal gyrus (STG, including Wernicke's area), middle temporal pole (mATP), superior temporal pole (sATP), inferior frontal gyrus pars opercularis (IFG_oper), inferior frontal gyrus pars triangularis (IFG_tri), and inferior frontal gyrus pars orbitalis (IFG_orb). Regions are defined following the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) as implemented in SPM12. I then computed the average BOLD signal for each of the 24 participants and six ROIs.

For each of the six ROIs, I fitted three linear mixed-effect models predicting the average BOLD signal. The first model (Base, 5 below) contains as predictors only probabilistic information (lexical frequency and surprisal) relative to each word. Estimates from the motion-correction algorithm (three rotations and three translations per run) were additionally added as regressors of no interest. In order to assess the effect of dependency and phrase structure measures to ROI's BOLD signal, models 6 and 7 were fitted with one of the syntactic measures (DSlrels, PSxps) each in addition to the same covariates of the Base model. In addition, I included by-subject random intercepts, as well as the by-subject random slopes for surprisal and log-transformed word frequency.

(5) Base model : $BOLD\ 1 + lexfreq + surprisal + m1 + m2 + m3 + m4 + m5 + m6 + (1|subject) + (1 + surprisal|subject) + (1 + lexfreq|subject)$

(6) DSlrels model : $BOLD\ 1 + lexfreq + surprisal + DSlrels + m1 + m2 + m3 + m4 + m5 + m6 + (1|subject) + (1 + surprisal|subject) + (1 + lexfreq|subject)$

(7) PSxps model : $BOLD\ 1 + lexfreq + surprisal + PSxps + m1 + m2 + m3 + m4 + m5 + m6 + (1|subject) + (1 + surprisal|subject) + (1 + lexfreq|subject)$

I compare the syntactic models (DSrels and PSxps model) against the Base model in order to test whether the introduction of the syntactic measure significantly improves the fit to the data. I also directly compare the DSrels and the PSxps models in order to test for specific syntactic structure selectivity in the six ROIs. Model comparisons are performed using the likelihood-ratio test. The models are fit by maximum likelihood.

4.5.2 PPI analysis

The ROI analysis introduced above is aimed at determining the contribution of the structural measures to the activity of left inferior frontal, superior temporal, and antero-temporal regions of the brain: areas that are claimed to be responsible for structural analysis of linguistic stimuli. In order to investigate the interplay between these (and other) brain regions, I additionally introduced a PPI analysis. PPI is a brain connectivity analysis method, developed to estimate context-dependent changes in functional connectivity cortical areas (K. J. Friston et al., 1997; K. J. Friston, 2011). It models the way brain activity is determined by the activity of a preselected seed region when modulated by experimental conditions or parameters (modulator). The analysis takes the activity of the seed region (**physiological component**) and a modulator (**psychological component**) and fits a voxel-wise linear model using as predictor of interest the product of these two components (**psychophysiological interaction**). In this way, the PPI identifies brain regions whose activity depends on an interaction between psychological context (the task) and physiological state (the time course of brain activity) of the seed region (O'Reilly et al., 2012).

The activity of each seed region was computed by fitting a general linear model containing as predictors the structural measures and as covariates lexical frequency and surprisal and parametric head movement. The eigenvalue of the voxels inside the ROI showing supra-threshold activation for the regressor of interest was used to compute the **physiological component** of the PPI. This was conducted at single subject-level with a significance level of $p < 0.05$. The regressor of interest used for ROI-wise voxel selection also acted as **psychological modulator** for the subsequent PPI analysis proper, which consisted of fitting another subject-level whole-brain general linear model using as regressor of interest the product of the seed activity and the modulator measures, and as covariates the seed activity and modulator themselves. The goal is then to identify those voxels (both at single subject and group level) that respond significantly to the interaction between seed activity and modulator.

4.5.3 Whole brain analysis

At the single-subject level, the observed BOLD time course in each voxel is subjected to a regression analysis, testing for voxels in which the covariates of interest (DSlrels, PSxps) explain a significant proportion of variance of that voxel's time course (K. Friston, 1995). Before the actual analysis, one regressor modeling the duration of each single word was created for each story. This regressor was convolved with the hemodynamic response function to account for the delay in BOLD activation respective to stimulus presentation. Besides the covariates of interest, log-transformed lexical frequency per-word – computed using the SubtlexNL corpus (Keuleers et al., 2010) – and per-word surprisal were introduced. They were used as regressor of no interest to statistically factor out effects of stochastic properties of the words. The estimates from the motion correction algorithm (three rotations and three translations per run) were additionally added as regressors of no interest.

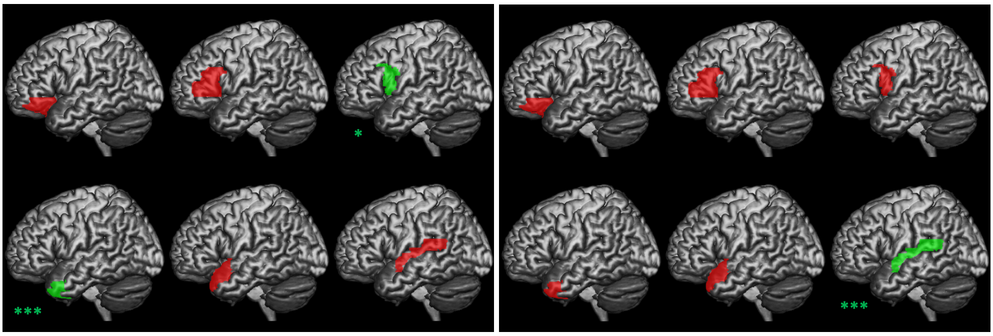
I am interested in assessing which voxels are more sensitive to DPlrels as compared to CPxps and vice-versa. In order to do so, I contrasted these two regressors of interest in order to identify voxels that are selective for each one of the regressors over and above the contribution of the other (DSlrels >PSxps, PSxps >DSlrels). The significance of these contrasts was assessed by computing the t-statistic over participants of this difference score for each voxel in the brain. The resulting multiple comparison problem was solved by means of combining a $P < 0.05$ voxel threshold with a cluster extent threshold determined by means of 1,000 Monte Carlo simulations, after the estimation of the smoothness of the data ((Slotnick et al., 2003)) applied for each separate contrast both for the single and the total models. Clusters of size exceeding the number of voxel threshold corresponded to statistically significant effects ($P < 0.05$ level, corrected for multiple comparisons).

4.6 Results

4.6.1 ROI analysis

Comparison against the Base model

I computed the likelihood-ratio test for the difference in fit between the Base model and each of the two Full models above across the 6 ROI's. This allowed me to test whether the introduction of syntactic measures significantly improves the fit of the linear mixed effect model to the BOLD signal. Table 4.3 reports the results of these analyses.



a: Dependency structure **b:** Phrase structure

Figure 4.4: Cortical illustration of the likelihood ratio test between the base model and DSlrels model (a), and between the base model and PSxps model (b). Green indicates ROIs where the likelihood ratio test returns significant results (*** p <0.001, ** p <0.01, * p <0.05). Red instead indicates ROIs for which the syntactic measure does not significantly improve fit. Note the different selectivity between anterior and posterior temporal areas of the network with regard to the type of syntactic structure.

The results indicate that there is a strong effect of PSxps in left STG (41.84, p <0.001). On the other hand, the DG structure building measure – DSlrels – has a strong effect in left middle TP (12.75, p <0.001) and a weaker effect on left IFG pars opercularis (3.83, p <0.05).

Comparison between models

Table 4.4 contains the results of the log-likelihood test between the DSlrels and PSxps models. Syntactic operation measures DSlrels and PSxps were directly compared to assess the prominence of one or the other as predictor of activity inside the ROI pool.

Left STG confirms a strong preference for phrase-structure parse measure with the model fitted with PSxps significantly outperforming a model fitted with DSlrels. DSlrels instead outperforms its PSG counterpart in middle ATP and in the pars opercularis and triangularis of the IFG.

ROI	DSlrels	PSxps
STG	1.10 (0.29)	**41.84 (0.00)*****
mATP	**12.75 (0.00)*****	0.77 (0.37)
sATP	2.04 (0.15)	0.25 (0.61)
IFG_oper	**3.83 (0.05)***	0.89 (0.34)
IFG_tri	2.90 (0.08)	0.30 (0.58)
IFG_orb	0.79 (0.37)	0.36 (0.54)

Table 4.3: Likelihood Ratio Test between Base model and each of the 2 models fitted with one of 2 syntactic measures derived either from dependency and phrase-structure parses.

Mod. comparison	**ROIs**		
PSxps >DSlrels	STG	**40.71**	**0.00*** **
DSlrels >PSxps	mATP	**11.97**	**0.009** **
	IFG_oper	**2.93**	**0.05***
	IFG_tri	**2.59**	**0.05***

Table 4.4: Likelihood Ratio Test between Dependency relations and phrase-structure phrase structures.

4.6.2 PPI analysis

The results presented in the previous section highlight a preferential selectivity for dependency structure in the left ATP and IFG, and a selectivity for phrase structure in the left superior temporal gyrus.

In this section, I present the results of a PPI analysis aimed at assessing the relation between activity in the ROIs as modulated by the processing of either phrase-structure or dependency structure. Since STG showed selectivity for phrase structures, and IFG and ATP for dependency structures, I conducted three separate whole-brain PPI analyses. I first checked for brain areas whose activity is driven by STG activity (physiological seed) modulated by PSxps structural measure (psychological modulator). I then used the activity of either the left IFG or the left ATP as physiological seeds, and DSlrels as modulator in order to assess the contribution of these areas and structure to the activity of the rest of the brain during language processing.

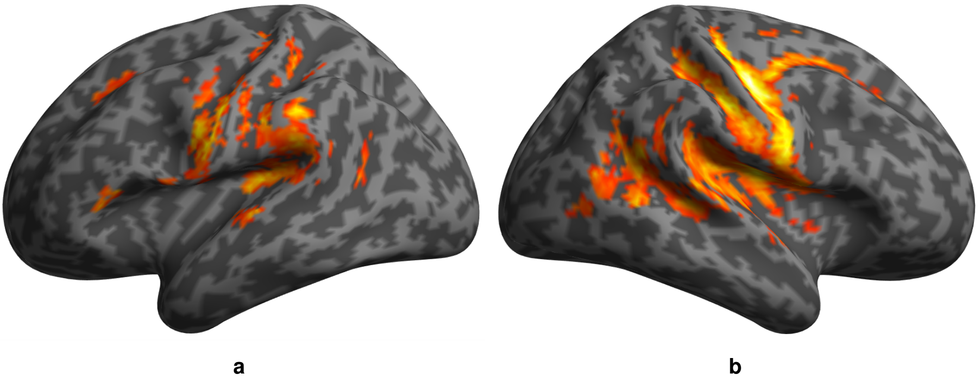


Figure 4.5: Results of the PPI analysis using as seed the left STG and as modulator PSxps.

Table 4.5 and Figure 4.5 reports the results of the PPI analysis using STG as physiological seed and PSxps as activity modulator. The results highlight large clusters in the bilateral central sulci (CS) and precentral gyri (PCG) encompassing both bilateral primary motor and premotor cortices. Activation is also observed for the bilateral posterior temporal and perisylvian cortices. Interestingly, activity in the left IFG is also driven by the interaction between the activity in the left STG and the PSxps measure.

Tables 4.6 and 4.7, and Figures 4.6 and 4.7 instead report the results of using DPlrels as modulator and ATP and IFG as physiological seeds respectively. These results indicate

Area	MNI coord.	T	p	cluster size
left IFG (triangularis)	-50 22 0	3.08	0.003	94
left PCG (Premotor)	-58 -4 28	3.16	0.002	2922
left MTG (Auditory)	-34 -30 10	3.12	0.003	
left CS (Primary motor)	-54 -12 36	3.00	0.003	
right CS (Primary motor)	36 -14 38	4.20	0.000	10896
right PCG (Premotor)	36 -12 45	3.62	0.001	

Table 4.5: Results of the PPI analysis using as seed the left STG and as modulator PSxps.

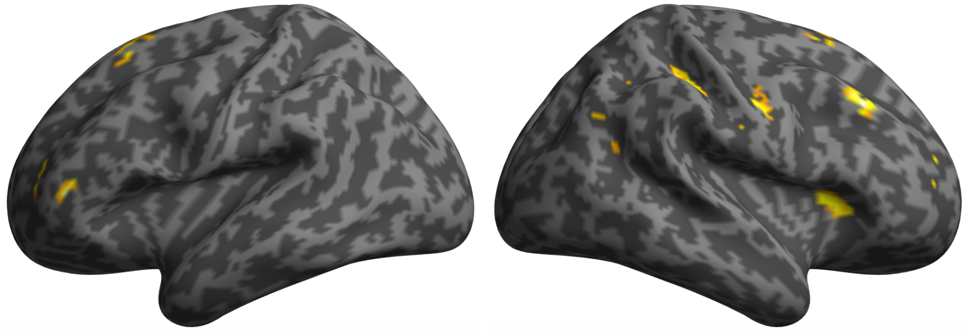


Figure 4.6: Results of the PPI analysis using as seed the left ATP and as modulator DPlrels

Area	MNI coord.	T	p	cluster size
left dorsolateral PFC	-24 0 20	3.27	0.002	3640
left Caudate	-28 36 16	2.70	0.007	224
	-30 42 6	2.69	0.007	
rigth dorsolateral PFC	44 26 40	3.34	0.002	142

Table 4.6: Results of the PPI analysis using as seed the left ATP and as modulator DPlrels

that the activity of the left ATP, modulated by DSlrels, explain the activity in a limited set of clusters located in the bilateral prefrontal cortex.

Instead, activity in the left IFG modulated by DSlrels explains the activity in the left supramarginal (SMG) and angular (AG) gyri in the posterior perisylvian cortex. Activity in the bilateral anterior prefrontal cortex (PFC) is also driven by this interaction, as is part of the middle banks of the right STG.

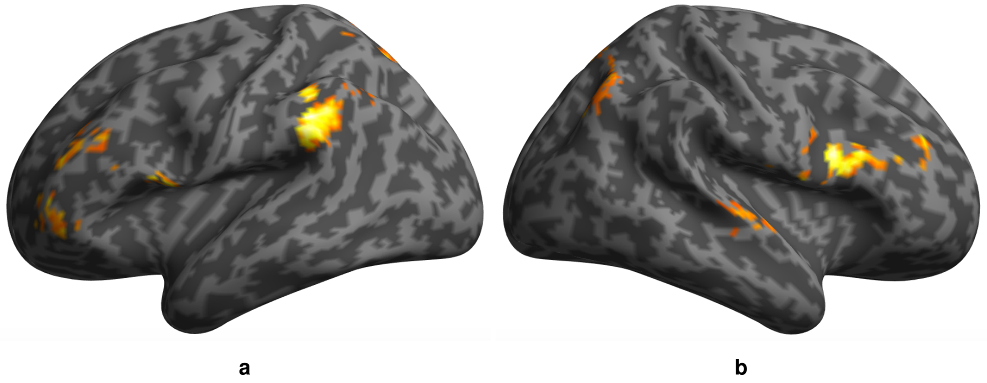


Figure 4.7: Results of the PPI analysis using as seed the left IFG and as modulator DPlrels

Area	MNI coord.	T	*p*	cluster size
left IFG (Triangularis)	-58 16 6	3.64	0.001	100
left SMG	-54 -40 34	3.71	0.001	625
left AG	-64 -42 26	3.18	0.002	
left anterior PFC	-44 44 20	3.36	0.001	484
right anterior PFC	44 44 22	3.72	0.001	1014
right SPG	30 -66 62	3.29	0.002	512

Table 4.7: Results of the PPI analysis using as seed the left IFG and as modulator DPlrels

4.6.3 Whole brain analysis

Tables 4.8 and 4.9 contain the results of the whole brain analysis contrasting PSG and DG measures.

The contrast between PSxps and DSlrels (PSxps >DSlrels, 4.8) highlights the role of IFG (orbitalis), AG, Fusiform gyrus and Hippocampus on the left hemisphere, and of the bilateral occipital cortex.

On the other hand, the comparison between DSlrels and PSxps (DSlrels >PSxps, 4.9) points to an involvement of the left superior frontal regions (SFG) and the right Caudate.

Area	MNI coord.	T	*p*	cluster size
left Hippocampus	-32 -14 -18	5.32	0.000	701
left MTG	-34 8 -22	3.36	0.001	
left Fusiform	-52 -42 -14	4.14	0.000	83
left Occipital	-12 -90 -10	3.88	0.000	1637
left AG	-30 -72 38	3.78	0.001	
left IFG (Orbitalis)	-20 30 -10	3.58	0.001	198
right Occipital	42 -80 8	3.48	0.001	2046

Table 4.8: Whole-brain results areas more sensitive to PSxps as compared to DSlrels (PSxps >DSlrels).

Area	MNI coord.	T	*p*	cluster size
left AG	-40 -58 22	5.00	0.000	9207
right Caudate	6 14 8	4.56	0.000	
left SFG	-12 52 26	3.12	0.002	216
right posterior CC	12 -22 38	3.32	0.001	139

Table 4.9: Whole-brain results areas more sensitive to DSlrels as compared to PSxps (DSlrels >PSxps).

4.7 Discussion

The goal of the experiment was to investigate whether parts of the brain, which have previously been implicated in syntax, are sensitive to different kinds of syntactic operations necessary to parse sentences. I have investigated if the brain activity of 6 left-hemispheric regions was better explained by dependency or by phrase structures (Section 4.6.1). The two grammars were not meant to represent contrasting hypotheses, and my aim was not to prove that one is a better formalism than the others. The present results suggests that both grammars can explain variance in cortical areas supposedly involved in syntactic processing, and that they appear to do so for different areas of the brain.

4.7.1 A syntactic division of labor

The results of a series of ROI analyses (Section 4.6.1) show that dependency structure measures significantly explain activity in left ATP and left IFG (opercularis), and that phrase-structure measures seem instead to explain activity in the left STG. These results overall seem to point towards a general division of labor between anterior temporal and inferior frontal areas – responsible for the computation of dependency representations of the sentence – and more posterior temporal areas involved, instead, in phrase-structure analysis.

These results partially differ from previous studies adopting phrase-structures as their formalism of choice to describe natural language syntax (Brennan et al., 2016; Nelson et al., 2017). Nonetheless these observations are more in line with studies such as Westerlund et al., 2015 that show ATP sensitivity to a wide range of compositional structures, including verb-argument and preposition-argument pairs. These results corroborate the idea that ATP – among the other regions studied here – works as a hub for sentential-level semantic composition where words are combined according to the argument structure of the sentences as captured by its dependency parse.

4.7.2 The role of the anterior temporal lobe in syntactic processing

The results indicate a relation between ATP activity and the number of left-hand side dependency relations at word level, which was adopted as a quantification of the operation the human brain is supposed to carry out in order to integrate each word in the dependency parse of the sentence it is embedded in (Section 4.6.1). This, therefore, seems to indicate that the ATP (together with parts of the IFG) acts as a combinatorial hub that binds together words according to relations similar to the ones characterized by a dependency grammar. These results are in line with previous studies that describe this area as a hub for composition during sentence comprehension (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2006, 2007).

Binary relations making up dependency structures resemble the 2-word stimuli that were found eliciting activation in the ATP by Baron et al., 2010 and by Westerlund et al., 2015. Baron et al., 2010 observed a modulation of activity in this region when adjective–noun pairs where presented. The relations holding between nouns are directly captured by the modification relations in the type of dependency parses I adopted in this study. Similarly, an interesting parallelism between the typified relations that constitute DG and the types of 2-word stimuli that were presented in Westerlund et al., 2015 might help understand why dependency structures might be a correct way of characterizing the type of job performed by the ATL. Dependency relations directly link pairs of words according to the type of role they play in a syntacto-semantic relationship. They can be grouped – broadly speaking – in verb-argument, i.e., the relation that is established between a predicate and its subject, object, or complement, or modifier typologies (other types exist, but for the sake of brevity I will not discuss them here). Modifier-type dependency relations can be exemplified by the link between an adjective and a noun, an adverb and a verb, a determiner and a noun, and so on. Westerlund et al., 2015 demonstrated that a wide range of "composition modes" affect the activity of the left ATP. These "modes" consist of 2-word sequences classified either as modification (Adjective-Noun Adverb-Verb, Adverb-Adjective) or argument saturation (Verb-Noun Preposition-Noun Determiner-Noun). These "modes" resemble the different types in which the dependency relations are classified. Therefore it seems natural to suggest that, on the basis of the results and the results found in the literature, dependency grammar offers a reasonable formalization of the type of structure employed (or constructed) in the left ATP.

4.7.3 The role of the left inferior frontal gyrus

The ROI analyses (Section 4.6.1) show that dependency-structure measures can explain activity also in left IFG (pars opercularis). This area plays center stage in several studies on language and syntactic processing and it is often associated with the activity in the left pSTG (Caramazza & Zurif, 1976; Friederici et al., 2005; Tyler et al., 2008; Snijders et al., 2008; Pallier et al., 2011; Zaccarella & Friederici, 2015; Zaccarella et al., 2015). The present results seem to indicate that this area might work in concert with the left ATP (Section 4.7.2) in building sentence-level representations that follow the structure described by word-word dependency relations. This is potentially compatible with the Memory, Unification and Control framework Hagoort, 2013b which predicates a role for IFG in integrating words into their sentential and discourse context.

Besides the ROI analyses, the PPI (Section 4.6.2) and whole-brain analyses (Section 4.6.3) provide a somewhat more complex picture. Activity in part of the left pars triangularis is explained by the activity of the left STG modulated by the phrase-structure, whereas activity in another portion of the same sub-region is linked to the activity in the pars opercularis modulated by the dependency structure. The whole-brain analyses instead indicate an involvement of the left pars orbitalis in phrase-structure processing.

In light of these observations, it is possible that different sub-regions of left IFG support the analysis of different syntactic structures, in concert with either the left ATP and the left STG. In particular, the pars opercularis might work in concert with the left ATP in building sentence-level dependency representations, whereas the pars orbitalis performs operations related to the ones carried on in the left STG and having to do with hierarchical phrasal representations of the sentence. The pars triangularis, in different ways linked to the computation of dependency structure in the ATP and the phrasal analysis in the STG, might – it is very tentative to say – act as a buffer between these two areas and their syntactic operations.

These explanations are still at the level of speculation, and I defer to further investigations before drawing any stronger conclusion.

4.7.4 Additional areas involved in dependency parsing

Whole-brain analyses highlighted an additional set of regions that are more sensitive to dependency structure as compared to phrase-structure measures. In addition to the ROI results, I observed the involvement of other brain structures: left AG, right posterior cingulate cortex, and left superior frontal gyrus.

These patterns of activation might indicate that dependency structures correlate with working memory mechanisms subserving syntactic parsing. PFC has been flagged out as a central player in working memory studies, including in the domain of language processing and sentence comprehension (D'Esposito & Postle, 2007; Nee & D'Esposito, 2016). In addition, Bonhage et al., 2014 reported the involvement of also the inferior parietal cortex (including AG) and areas bordering the cingulate cortex and the precuneus during the encoding in working memory of short sentence fragments (4 or 6 words).

One can conclude that the number of left-hand side relations (DSlrels) governing each word in the stimulus is simply modeling the load on working memory resources required for word-by-word sentence processing. In other words, given a sentence, the brain has to store in memory each word incrementally until the recipient of a dependency relation with those words is presented, and eventually integrated. In this sense, DSlrels only capture the number of words to keep in mind until a suitable dependent or head is read or heard.

Nonetheless, this interpretation does not seem to explain the whole picture with regard to dependency structure processing. As pointed out in Sections 4.6.1 and 4.7.2 above, there is a significant relation between DSlrels and the left ATP, an area that is not traditionally considered part of the working memory network. Therefore, rather than interpreting these results as suggesting that dependency measures simply capture working memory loads imposed by the number of words to integrate in the parse, it might be more accurate to claim that, while dependency-related activity in the left ATP indeed computes sentence-level structural analyses, activity in areas such as the cingulate, frontal and inferior parietal cortex might be well explained in terms of working memory support to the activity in anterior temporal and inferior frontal regions. Further work is required to shed light on this possibility.

4.7.5 Interaction between areas

The PPI analyses (Sections 4.6.2) were conducted in order to see what type of interaction exists between these syntax processing areas (left ATP, IFG, and STG) and the rest of the brain.

The fact that activity in the left STG, modulated by PSxps, seems to drive the activation in a small portion of the left IFG, besides a large network of bilateral central and precentral regions, might indicate that there is an interaction between phrase structure and dependency structure processing areas. This might also be supported by the almost reversed observation that IFG activity, modulated by DSlrels, explains activity in the left posterior perisylvian cortex (AG), and in a small portion of the right middle STG. Nonetheless, these results cannot allow me to strongly claim a causal interaction between these sets of areas.

4.8 Conclusions

In this paper, I investigated whether different brain regions are sensitive to different kinds of syntactic operations. In order to do so, I assessed dependency and phrase-structure descriptors of sentences as predictors of brain activity in the left anterior temporal pole, the left inferior frontal gyrus, and the left superior temporal gyrus – areas engaged during language processing.

I found that activity in the left ATP is better explained by DG measures as compared to PSG ones. These results differ from those reported in neuroimaging studies using only phrase-structures to describe syntax (Nelson et al., 2017). My results are related to the ones presented by Brennan et al., 2016. They predicted fMRI data in both the left ATP and left pSTG during narrative listening using syntactic metrics derived from Minimalist Grammars (MG) (Stabler, 1997), which derive syntactic descriptions combining both phrase-structure and dependency information. In a more recent study Li & Hale, 2019 showed that Brennan and colleague’s observed effect of using MG are still present in the left pSTG even after controlling for memory-based metric “structural distance" and a distributional-semantic metric indicating “conceptual combination". Nonetheless, two aspects distinguish my study with the ones by Brennan et al., 2016 and Li & Hale, 2019. First, in the present study I deliberately decided to keep phrase-structure and dependency measures apart under the assumption that their structural differences might explain activity in different areas composing the language network in the brain. As a matter of fact, our results showed how the left ATP and the left STG are selective for one or the other. Second, our results, obtained on Dutch instead of English, may be taken to support the cross-linguistic validity of these observations, keeping in mind that Dutch and English are closely related.

The present observations regarding the role of the left ATP are also in line with studies such as Westerlund et al., 2015 that show how this area is sensitive to a wide range of compositional structures, including verb-argument and preposition-argument pairs. This seems to corroborate the role in sentential-level semantic composition for the ATP, and the validity of dependency grammar as a formalism describing sentential structures.

Moreover, a series of PPI analyses investigating the interaction between each of the ROI's and the rest of the brain show that the activity in the left ATP might be driven by the activity in the left STG. These results, while confirming a division of labor between brain regions, seem to point to an ancillary role of STG and phrase-structure building, subserving the dependency-style analysis that an area such as the ATP seems to perform.

5

Effects of Dependency Structure on Eye Regressions During Reading

After having shown in Chapter 4 that dependency and phrase-structure grammar might be processed in distinct areas in the brain, in this chapter, I focus only on the former, and show that further proofs of its validity as an approximation of syntactic processing during natural language comprehension can be obtained by analyzing the patterns of eye movements observed during reading. Backward saccades during reading have been hypothesized to be involved in structural reanalysis, or to be related to the level of difficulty posed by a sentence. In this chapter, I test the hypothesis that they are instead involved in online syntactic analysis. If this is the case, I expect that saccades will coincide, at least partially, with the edges of the relations computed by a dependency parser. In order to test this, I analyzed a large eyetracking dataset collected while 102 participants read three short narrative texts. My results show a relation between backward saccades and the syntactic structure of sentences.

This chapter is based on:

Lopopolo, Alessandro, Stefan L. Frank, Antal van den Bosch and Roel M. Willems. 2020. Dependency Parsing with your Eyes: Depend-ency Structure Predicts Eye Regressions During Reading. *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* 2019. NAACL, Minneapolis (MN) 2019.

In Chapter 4, I have shown that the cortical network responsible for the syntactic analysis of the linguistic input can be partitioned in sub-networks sensitive to two different types of grammars: phrase-structure (PSG), and dependency grammar (DG). These grammars therefore prove to be valid approximations of the syntactic analyses carried out by the brain. In this chapter, I will focus only on DG, and show that further proofs of its validity as an approximation of syntactic processing during natural language comprehension can be obtained by analyzing the patterns of eye movements observed during reading.

5.1 Theoretical background

5.1.1 The role of regressions in text comprehension

Regressions (backward saccades) are relatively rare, usually occurring only with 15 to 25% of the words (Rayner & Pollatsek, 1995). They do not seem to be random, however. Regressions typically aim at specific word locations, moving fixation from the current word back to one of the previously encountered words (Vitu, 2005). Nonetheless, their function in language comprehension is still debated. Here I will review two proposed explanations: the first links regressions to the difficulty of text processing; the second instead sees them as tools for language processing, not necessarily linked to processing difficulties or errors. According to the first proposal, regressions only start to play a role in reading once difficulties are encountered; according to the second proposal, they are part and parcel of regular reading.

Regression as a response to comprehension difficulty

The first hypothesis interprets regressions as part of the reanalysis of textual input due to encountered comprehension problems. In a milestone study, Altmann et al., 1992 introduced the notion of regression-contingent analysis, based on the assumption that regressive eye movements are a necessary consequence of subjects being garden-pathed. A garden-path effect occurs when readers incrementally construct an incorrect interpretation of a sentence as a consequence of its locally ambiguous syntactic structure. This does not necessarily mean that the presence of a difficult structure, leading for instance to the reader being garden-pathed, triggers a regression. Rayner and colleagues reported data showing that strong garden path effects can sometimes occur without triggering any regressions (Rayner & Sereno, 1994; Castelhano & Rayner, 2008). Nonetheless, other studies have given support to the idea that regressions are linked to textual ambiguity and contextual difficulties. Readers make more regressions when the text is complex (Rayner & Pollatsek, 1995), when the topic changes (Hyönä, 1995), when the text contains grammatical errors or ambiguities (Reichle et al., 2013), or when they encounter information that disambiguates the preceding text (Blanchard & Iran-Nejad, 1987; Frazier & Rayner, 1982). The general hypothesis holds that the probability of regressions and their span might depend on the difficulty of the text. Therefore these regressions might allow the reader to reread information that has been missed, forgotten, or wrongly interpreted (Rayner, 1998).

Regression as a tool for comprehension

The alternative explanation focuses instead on the role of eye movements as a tool in language processing, used independently from the structural difficulty of the input. The idea is that regressions help the reader reactivate cognitive information that is associated with the regressed-to location. Kennedy, 1992 refers to this as the Spatial Code Hypothesis. The hypothesis is that readers use the position of words on the page as a support to their working memory by reactivating previously read words associated with information relevant for the processing of the word from which the regression originated (O'Regan, 1992; Spivey et al., 2004). This hypothesis has been criticized by Booth & Weger, 2013. They presented three experiments showing that readers' comprehension is not hindered when reading conditions inhibit or discourage visual access to already read material. In their Experiment 1, readers knew that candidate targets for regression were no longer available for rereading. Experiment 2 discouraged regressions by forcing readers to follow a visual placeholder on the stimulus while it was also presented in auditory form. Finally, in Experiment 3, candidate targets for regression were manipulated after reading. In all these three experiments, readers showed no hindered comprehension of the presented stimulus sentences. As an entailment of these results, Booth and Weger suggested that readers do not use regressions to cue their memory for previously read words.

My hypothesis is that readers might make use of regressions to reactivate previously read information in the context of naturalistic language comprehension, in order to help compute linguistic information.

I want to examine whether there is an alignment between patterns of regressions and word-to-word syntactic relations, as described by the dependency structure of the stimulus. I hypothesize that regressions play a role in syntactic parsing that may go beyond the reanalysis of ambiguous material. I do not deny their role in reanalysis and repair, but I rather stand with the hypothesis that they allow rereading and cueing of previous words, as an aid to memory, when this is required for a successful construction of a syntactic representation of the text.

In order to test this hypothesis, I rely on an eyetracking dataset that was collected during normal text reading of unmodified literary narratives. I assess whether there is a relation between the number of eye regressions from the words and the number of syntactic relations that those words entertain with their preceding text. These syntactic relations are derived from the dependency structures (described in Section 5.3) of the sentences composing the stimuli of the eyetracking dataset.

5.1.2 Regressions and sequence processing

The hypothesized relation between dependency structure and eye movement taps into a broader debate on whether language processing relies mainly on the sequential structure of the input or whether it involves the computation of non-sequential syntactic parses (Jackendoff & Wittenberg, 2014). Undeniably, the linguistic stimulus is presented as a string of symbols; nonetheless, regressions seem to counter the notion that it is processed strictly in a sequential order. If these eye movements are involved only in reanalysis, then their existence does not necessarily contradict sequential processing accounts. They can be explained as an ”emergency recovery” operation that takes place only in cases of processing difficulties. On the other hand, if I find evidence of a relation between saccades and syntactic dependency structures independent from processing difficulty, then I might conclude that saccades offer behavioral evidence that processing involves the computation of non-sequential structures.

This question is related to the line of research in psycholinguistics and neuroscience investigating the computation of syntactic structures during language processing. In this context, sequential structures are usually contrasted with hierarchical ones, where input items are grouped into larger units, which in turn are (possibly recursively) grouped in even larger units. These larger units are commonly referred to as syntactic constituents or phrases and have a central position in theoretical linguistics (Chomsky, 1965; Jackendoff, 2003, 2007). An increasing amount of evidence against a strictly hierarchical processing of language has emerged over the past decades. Psycholinguistic studies have supplied evidence suggesting that the mere sequential properties of the stimulus are sufficient to explain aspects of human behavior during reading and listening. Frank et al., 2012 provide a review of evidence from cognitive neuroscience, psycholinguistics, and computational modeling studies supporting the hypothesis that hierarchical structure may not play a central role in language processing and acquisition and that sequential structure instead has significant explanatory power. They argue that hierarchical structure is rarely needed to explain behavioral and neural correlates of language processing in vivo. In contrast with these findings, recent neuroimaging studies have delineated a slightly more complex landscape in which both hierarchical and sequential processing may be carried out simultaneously by the human brain during language processing (Brennan et al., 2016; Nelson et al., 2017).

Dependency parses are different from constituency parses as they lack the non-terminal nodes characteristic of constituency parses. Nonetheless, they still constitute a non-sequential type of structure. Demonstrating a relationship between eye movement and such structure will provide evidence for the non-sequentiality of language processing, at least in the context of text reading.

5.2 Related work

The present work studies the relation between eye movements during reading and the dependency structure as produced by a dependency parser (see Section 5.4.1 for more details). Several other studies tested language processing hypotheses by using computational models as predictors of eye movements during sentence reading.

Boston et al., 2008 demonstrates the importance of including parsing costs implemented as surprisal as a predictor of comprehension difficulty in models of reading. They showed that surprisal of grammatical structures has an effect on fixation durations and regression probabilities.

Demberg & Keller, 2008 compared linguistic integration cost computed as a function of dependency relations distances and word surprisal as predictors of gaze duration. They showed that distance is not a significant predictor of reading times except for nouns. On the other hand, they demonstrate that surprisal can predict reading times for arbitrary words in the corpus, concluding that the two predictors may capture distinct aspects of naturalistic language processing.

In the context of Natural Language Processing, Klerke et al., 2015 used eyetracking data as a metric for the quality of automatic text simplification and compression, which are operations used in machine translation and automatic summarization. Their proposal is grounded in the hypothesis that eye movements are related to perceived text difficulty (Rayner & Pollatsek, 1995), one of the two hypotheses I have introduced in Section 5.1 above.

5.3 Dependency structure

As already explained in Section 4.3.2, dependency grammar describes a sentence as a set of relations between words (heads) and their dependents. These relations are called dependencies and correspond to grammatical functions and – together with the words they link – are the only descriptive elements composing the structure, which has the form and properties of a directed graph (Tesnière et al., 2015; Mel’čuk, 1988; Nivre & Kübler, 2009).

This type of structure lacks phrasal non-terminal constituents. In addition, it is not strictly sequential, or put differently, it is not isomorphic to the sequence of items that makes up the stimulus. This is based on the fact that the dependency relations can hold between words that are non-consecutive or possibly even far apart in the sentence.

The assumption is that during reading, these links are created once a suitable candidate for the second term of the dependency is introduced. Therefore, online dependency parsing proceeds by introducing one word at a time, and by looking back at the prefix in order to assess whether this novel input is a suitable candidate for a dependency link with a preceding word that has not yet been matched.

5.4 Materials and methods

For the purpose of this study, I have used the eyetracking narrative dataset described in Chapter 2, Section 2.2. The dataset consists of data of 102 native Dutch speakers collected while the participants were asked to read three Dutch narrative texts. The eyetracker data used in this study was originally collected for a study on mental stimulation during literary reading by Mak & Willems, 2018 at Radboud University, Nijmegen, the Netherlands.

Eyetracker measures

For the present study, I focus on the number of eye regressions. A regression consists of a fast eye movement from a word back to a previous word.

5.4.1 Dependency parsing

The texts of the three stories presented to the participants were fed to the ALPINO toolbox for Dutch natural language processing (Noord, 2006) to generate a dependency parse for each of their sentences. The parser creates a structure composed of dependency triples consisting of a head word, the type of dependency relation, and its dependent word. A parse is produced for each sentence independently; therefore, no relation can be assigned between words belonging to different sentences. The output of the parser was manually checked in order to prevent tokenization and sentence segmentation errors.

Number of dependency relations

As described in Section 5.3, every word in a sentence entertains at least one relationship with another word in the same sentence. Every non-final and non-initial word can have relations with a variable number of other words on its right and its left. Because I are interested in eye regressions, I decided to focus my attention only on relations between a word w and its

	the	man	saw	a	brown	dog	in	the	park
N_head	0	0	0	0	0	1	1	0	1
N_dependents	0	1	1	0	0	2	0	0	1

Table 5.1: Number of dependency relations per word w that correspond to words in w's own preceding context.

preceding context. Therefore only relations with a head and possible dependents on the words preceding w are counted.

From the dependency structure of a sentence, I derived the following count measures:

- ***N_head*** indicates the presence of a syntactic relation between w_i and a word in $w_{1:i-1}$ that is head of w_i;
- ***N_dependents*** counts the number of syntactic relations between w_i and words in $w_{1:i-1}$ that are dependents of w_i.[1]

Measure *N_head* is a binary variable indicating whether word w has a head in its left context $w_{1:t-1}$. This is because every word has one and only one head.

For example, the word *dog* in Sentence 1 from Chapter 1 (*The man saw a brown dog in the park*) has one head relation with a word on its left (*saw*), two dependents on its left (*a* and *brown*), and none on its right. On the other hand, the word *park*, being sentence-final, does not have any links on its right, but it has one head (*in*) and one dependent (*the*) on its left. Table 5.1 contains the count measures (*N_head* and *N_dependents*) for this sentence.

5.4.2 Descriptors not related to dependencies

I am interested in the effect of syntactic structure, implemented as dependency relations, on the pattern of regressions. For this reason, it is necessary to control for other possible quantifiable factors affecting these eye movements. I chose to use log-transformed lexical frequency and surprisal.

Base-2 log-transformed lexical frequency per word was computed using the Subtlex NL corpus (Keuleers et al., 2010). Surprisal was computed from a second-order Markov model, also known as trigram model, trained on a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the Dutch section of Corpora from the Web (Schäfer & Bildhauer, 2012). Surprisal of word w_t is the negative logarithm of the conditional probability of encountering w_t after having read sequence w_{t-2}, w_{t-1}, or:

---

[1] Note the similarity with the DSlrels measures of Chapter 4.

$-\log P(w_t|w_{t-2}, w_{t-1})$. The computation was performed by the SRILM toolbox (Stolcke, 2002).

Frequency and surprisal are computed in order to control for processing difficulties. Intuitively, infrequent words and words with high surprisal are more difficult to retrieve (and possibly to integrate) with their preceding context. Controlling for processing difficulty is motivated by the alternative hypothesis regarding the role of regressions as depending on the level of complexity posed by a linguistic input.

In addition to frequency and surprisal, I also use word position in the sentence. Intuitively, the probability of regressing from a word to its previous context increases linearly with the position of the word in a sequence. By controlling for it, I ensure that the eye movements are not due simply to the opportunity given by the larger target pool to regress to.

5.4.3 Analyses

I fitted two logistic mixed-effect models predicting eye regressions. The first model (**null**, Eq. (5.1a) below) contains as predictors only the position of the words in their sentences (*word_order*), and probabilistic information consisting of the above-mentioned log-transformed frequency (*freq*) and surprisal (*surp*). The second model (**full**, Eq. (5.1b) below) contains as predictors of interest also the number of left-hand side dependency relations (i.e. *N_head* and *N_dependents*) of each word. In addition, I included by-participant and by-word random intercepts, as well as by-participant random slopes for *word_order* in the **null** model and for *word_order*, *N_head* and *N_dependents* in the **full** models.

I expect the model's fit to improve significantly after the inclusion of the measures derived from the dependency parse as regressors. The increase in model fit is quantified by the χ^2-statistic of a likelihood-ratio test for significance between the **null** and **full** models and is taken as the measure of the fit of *N_head* and *N_dependents* measures at each word to the probability of a regression being generated at each word.

$$\begin{aligned}\textbf{null} &: \textit{eye\_regressions} \\ &\sim \textit{word\_order} + \textit{surp} + \textit{freq} + (1|\textit{word}) + (1 + \textit{word\_order}|\textit{participant})\end{aligned} \tag{5.1a}$$

$$\begin{aligned}\textbf{full} &: \textit{eye\_regressions} \\ &\sim \textit{word\_order} + \textit{surp} + \textit{freq} + \textit{N\_head} + \textit{N\_dependents} \\ &\quad + (1|\textit{word}) + (1 + \textit{word\_order} + \textit{N\_head} + \textit{N\_dependents}|\textit{participant})\end{aligned} \tag{5.1b}$$

The models are fit by maximum likelihood (Laplace Approximation) and with a binomial distribution.

5.5 Results

5.5.1 Regression model analysis

In the results below I first describe the fit of each of the two models (**null** and **full**) separately, then I report the results of the model comparison analysis using the χ^2-statistic.

Table 5.2 presents the fitted **null** model. Table 5.3 shows the fitted **full** model. The head and dependent regressors have significant effects on the number of regressions (*eye_-regressions*) - (N_head: $\beta = 0.242, p < .0001$; N_dependents: $\beta = 0.046, p < .0005$).

In addition, both word frequency (*freq*) and surprisal (*surp*) have a significant negative effect. The negative effect of frequency might be due to less frequent words being more difficult to retrieve from memory, therefore triggering a regression to gather more contextual information to help word processing. The negative effect of surprisal indicates that the larger the surprisal of a word – therefore more difficult its integration into the context – the less probable the reader is to regress to the word's previous context. Mak & Willems, 2018 reported a positive effect of surprisal on the number of incoming saccades, that is, eye movements into a word back from subsequent parts of the text.

	Estimate	**Std. Error**	**z value**	**$\Pr(>\|z\|)$**
(Intercept)	-1.616	0.049	-32.9	$< .0001$
word_order	0.035	0.007	5.2	$< .0001$
surp	-0.140	0.013	-11.1	$< .0001$
freq	-0.165	0.028	-6.0	$< .0001$

Table 5.2: Fixed effects for the **null** model

	Estimate	**Std. Error**	**z value**	**$\Pr(>\|z\|)$**
(Intercept)	-1.798	0.049	-36.6	$< .0001$
word_order	0.019	0.006	3.0	$< .003$
surp	-0.102	0.012	-8.0	$< .0001$
freq	-0.125	0.027	-4.6	$< .0001$
N_head	0.242	0.016	14.8	$< .0001$
N_dependents	0.046	0.013	3.6	$< .0005$

Table 5.3: Fixed effects for the **full** model

model	**df**	**AIC**	**BIC**	**deviance**	χ^2	χ^2**df**	**Pr**
null	8	401023	401111	401007			
full	17	400302	400489	400268	738.87	9	$< .0001$

Table 5.4: Results of log-likelihood comparison between **null** and **full** model.

In order to test whether the introduction of head and dependent measures improves the fit of the logistic mixed effect model to eye regressions, I computed the χ^2-statistic of a likelihood-ratio test for the difference between the **null** and **full** models above. The χ^2 is taken as the measure of the fit of the dependency measures to the probability of a regression being initiated at each word. Table 5.4 reports the results of the test, showing the difference in model fit to be significant ($\chi^2 = 738.87$, $p < .0001$).

5.5.2 Analysis of regression counts

The results of the regression model comparison indicate that regressions are partially driven by the presence of left-hand side dependency relations. In order to corroborate these observations, I counted the number of times regressions generated from each word do actually land on preceding words that are heads or dependents of that word. As reported in Table 5.5, it turns out that of the 110,336 regressions, about 40% do actually land on a head or dependent of the words they originate from. These are referred to as matches. The analyses were limited only to regressions landing within sentence boundaries. In the table, "misses" refers to the regressions that land on targets that are neither head nor dependent of the word they originated from.

tot nr of regressions:	110336
tot nr of matches:	46378
tot nr of misses:	63958

Table 5.5: Total numbers of regressions, matches (i.e. regressions that land on heads or dependents), and misses (i.e. regressions do not land on heads or dependents of the word they originated from).

A χ^2-test of independence was performed to assess the relation between having a dependency relation with a word and generating a regression to that word. The test was computed independently for 10 separate left-hand side distances $d = [-10 : -1]$. In other words, for $d = -1$, I want to assess whether there is a relation between having a dependency relation with the preceding word and looking back at that word; for $d = -2$, I want to assess whether there is a relation between having a dependency relation with the preceding word at position -2 and looking back at that word, and so on for the other considered distances.

Table 5.6 contains the per-distance results of the χ^2 analyses. An association between presence of a dependency relation at position d and the generation of a regression to that position is significant for distances -1 ($\chi^2 = 132.52, p < 0.001$), -2 ($\chi^2 = 678.14, p < 0.001$), -3 ($\chi^2 = 8.05, p < 0.005$), and -4 ($\chi^2 = 13.68, p < 0.001$). For all other tested distances (between -5 and -10) the association was not significant (see Figure 5.1). For $d = \{-1, -2, -3, -4\}$:

- The fraction of words w_i in a dependency relation with w_{i-d} that originate a regression of length $-d$ is significantly higher than the fraction of w_i not in a dependency relation with w_{i-d} originating a regression of length $-d$;
- The fraction of words w_i with a regression of length $-d$ that are in a dependency relation with w_{i-d} is significantly higher than the fraction of w_i without a regression of length $-d$ that are in a dependency relation with w_{i-d}.

dist	+dp+reg	-dp+reg	+dp-reg	-dp-reg	χ^2
-1	29245	19520	305931	228111	**132.52****
-2	5711	14798	113028	449270	**678.14****
-3	937	6239	68785	506846	**8.05***
-4	309	2641	49324	530533	**13.68****
-5	76	1370	33153	548208	0.55
-6	35	850	29823	552099	2.69
-7	17	530	23728	558532	1.42
-8	13	355	17687	564752	0.29
-9	5	261	16774	565767	1.08
-10	3	263	13785	568756	2.18

Table 5.6: χ^2 analyses for distances $-10 : -1$. **+dp+reg** indicates the number of words in the corpus having a relation and a regression at $-d$; **+dp-reg** the number of words having a relation but not a regression at $-d$; **-dp+reg** number of words not having a relation but having a regression at $-d$; **-dp-reg** not having nor relation nor regression at $-d$ ($** = p < 0.001$, $* = p < 0.01$).

This seems to indicate that the effect of the dependency structure of a sentence on the pattern of outgoing eye-movements from a word is present only for short-distance relations (between a word and its four preceding words).

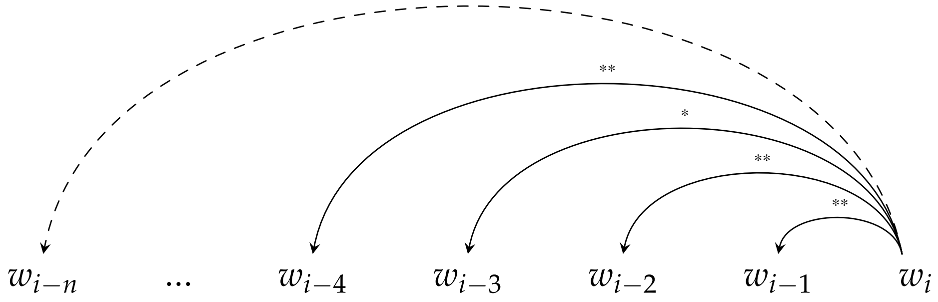


Figure 5.1: The effect of dependency relations on regressions from w_i is significant only for the preceding four words - further away saccades might not be influenced by a possible relation with w_i ($** = p < 0.001$, $* = p < 0.01$).

It is important to keep in mind however, that the number of dependency relations found by the parser is much higher than the actual number of matches. This is simply because the parser does assign at least a head to each and every word in the text (even words in isolation are assigned a 'root' head), whereas a regression is a relatively rare event (under normal conditions, using naturalistic language). The present work aims at demonstrating that regressions are related (also) to the structure of the dependency graph. The results I have produced so far point in that direction. In other words, it is possible to affirm that if a regression takes place, it might be triggered by the presence of a dependency relation between the word it is generated from and the word it lands on.

5.6 Conclusion

In this chapter, I have presented an analysis investigating whether the eye movements of readers may be influenced directly by the syntactic structure of the sentence. I tested this using shallow non-hierarchical structures computed by a dependency parser. The hypothesis was that the path of regressions from a word to an earlier word coincide, at least partially, with the edges of dependency relations between these words. I used a substantially large eyetracking dataset collected, while 102 participants were engaged in reading three short narrative texts.

The results of a mixed-effect regression analysis indicate that there is a significant effect of the number of left-hand side dependency relations on the number of backward saccades. This effect is well above chance even after correcting for word position in the sentence and word frequency and surprisal – measures held to explain a large part of natural language processing behavioral and neural correlates. These results are corroborated by the observation that about 40% of backward saccades do indeed land on target words engaged in dependency relations according to the syntactic structure of the sentences composing my stimuli.

The length of the regressions seems to be relatively short: the vast majority being shorter than three words, with a predominance of regressions one position backward. The results of a series of χ^2 tests for independence shows that there is a significant association between the presence of a dependency link and backward saccading between two words holds only for pairs, which are not further apart than four positions. This might indicate that the eye regressions are involved predominantly in dependency parsing at the local level, rather than at long distance.

Altogether these results converge on the idea that eye movements reflect, among other things, the shallow syntactic structure of language, corroborating the idea that humans do engage in online syntactic analysis of the input – at least in the form of local dependency parsing.

The evidence that eye gazes partially land onto previous words in a syntactic relation with the word they started from, even after controlling for complexity, seems to point towards an active role of regressions during syntactic parsing. In other words, the cognitive role of eye regressions emerging from this study seems to be in line with Kennedy's Spatial Code Hypothesis which posits that regressions help the reader reactivate cognitive information that is associated with the regressed-to location (Kennedy, 1992; O'Regan, 1992; Spivey et al., 2004).

6

Tracing Lexical Retrieval and Integration in the Brain

The meaning of a word depends on its lexical semantics and on the context in which it is embedded. Different brain networks carry out these different aspects of determining word meaning during comprehension, lexical retrieval and context integration, at different latencies. In this chapter, I investigate how lexical retrieval and integration are implemented in the brain by comparing MEG activity to word representations generated by computational language models. I test both non-contextualised embeddings, representing words independently from their context, and contextualised embeddings, which instead integrate contextual information in their representations. Using representational similarity analysis over cortical regions and over time, I observed that brain activity in the anterior temporal pole and inferior frontal regions between 300 and 500 ms after word presentation shows higher similarity with contextualised word embeddings compared to non-contextualized embeddings. Non-contextualised word embeddings, compared to contextualized embeddings, show higher similarity mainly in earlier latencies in the lateral and anterior temporal lobe – areas and latencies related to lexical retrieval. My results highlight how lexical retrieval and context integration can be tracked in the brain by using computational word embeddings. These results also suggest that the distinction between lexical retrieval and integration might be framed in terms of context-independent and contextualised representations.

This chapter is based on:

Lopopolo, Alessandro, Antal van den Bosch, Jan-Mathijs Schoffelen and Roel M. Willems. 2020. Words in context: tracking context-processing during language comprehension using computational language models and MEG.

In order to comprehend a sentence, before even computing its syntactic structure, it is necessary to retrieve the meaning of its composing elements – for instance the words that make it up – from the mental lexicon. Lexical retrieval is therefore one of the main basic operations underpinning language comprehension, the other is integration (Damasio et al., 1996; Caramazza, 1996; Jackendoff, 2003; Hagoort & Levelt, 2009; Friederici, 2011; Ullman, 2012).

In accordance with the discussion presented in Baggio & Hagoort, 2011, it is difficult to directly test the nature of the distinction between retrieval and integration using traditional task-based experimental paradigms, since it might be impossible to devise a task tackling only one operation while leaving the other untouched. Therefore, in the study presented here, I instead propose to use computational semantic models to further investigate the neural basis of these two basic operations. The models adopted in the present study are able to represent linguistic units as either independent or dependent from their context of occurrence. They are usually referred to in the computational linguistic literature as non-contextualized and contextualized embeddings (Mikolov et al., 2013; Peters et al., 2018). By comparing them to neural data collected during sentence comprehension, I aim to show that integration is approximated by contextualized embeddings and that it is a separate process (both functionally and physiologically) from lexical retrieval from memory.

Brain activity related to language comprehension is composed of processes that involve different areas of the brain at different moments in time following the onset of the stimulus (Friederici et al., 2000; Friederici, 2002; Hagoort, 2005). It is therefore capital to show that the putative similarity between a model and a brain process concerns not only areas associated with such process, but also that it does so in a time frame that is compatible with the time course of language processing. For this reason, I use a magnetoencephalographic (MEG) dataset collected during sentence reading. MEG records brain activity at the level of milliseconds, and with a reasonable anatomical resolution, making it ideal for a study interested in the when, and not only, the where of a specific neural process (Schoffelen et al., 2019).

6.1 Lexical processing in the brain

6.1.1 Neural loci

Semantic memory, a component of long-term memory, acts as the storage of knowledge and representations of basic linguistic units, such as words. In a simplified manner, memory can be defined as the mental lexicon, or the equivalent of a vocabulary in which the representations of words are stored and wait to be retrieved during production or comprehension (Hagoort, 2013a; Binder et al., 2009; Binder & Desai, 2011). The left temporal cortex, the middle temporal and inferior temporal gyri in particular, are considered central for the cortical implementation of semantic memory. In addition, also the anterior temporal pole (ATP) has been singled out as playing an important role. This has been confirmed by both studies on semantic dementia (Hodges et al., 1992; Rogers et al., 2004), and by a large neuroimaging literature (Tyler et al., 2004; Bright et al., 2005; Moss et al., 2004; Rogers et al., 2006). According to the hub and spoke model by Patterson et al., 2007, the ATP is supposed to produce supra-modal concept representations by collecting and controlling modality-specific features from sensorimotor areas.

Integration is a process that operates on representations retrieved from semantic memory and combines them in larger constructs. As explained in Chapter 1, I adopt a definition of integration based on the concept of context. In this light, a retrieved representation is integrated by embedding it in the linguistic context in which it happens to be uttered. Cortical areas responsible for this operation seem to belong to the inferior frontal gyrus and the perisylvian and temporal cortex (Hagoort, 2005, 2013a).

Integration also involves anterior temporal areas. A series of studies reported an increase in activity in these areas, in particular the ATP, during when the subjects were presented with sentences as compared to when they were presented with lists of words (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2006, 2007). The rationale behind this claim is that a structured linguistic input, such as a sentence, requires the integration of its constituent words, whereas a list of unconnected words does not. This claim is furthermore confirmed by the observation that the activity recorded in the ATP correlates with the presentation of word pairs in specific syntactic relations, such as subject-predicate, or adjective-noun (Baron et al., 2010; Baron & Osherson, 2011; Bemis & Pylkkänen, 2011, 2013; Westerlund et al., 2015).

6.1.2 Timing of processes

Besides the cortical loci of processing, sentence processing is characterized by a specific temporal profile that describes the timing of each of its sub-processes (Friederici, 2002; Humphries et al., 2007). The earlier stages mainly concern the recognition of the word from its auditory (for spoken words) or graphic (for written words) image and involve primary auditory or visual areas between the onset of a word and 150-200 ms. The phases that interest this analysis are the so-called Phase 1 and Phase 2, as described by Friederici, 2002.

Phase 1 takes places after the word form has been identified, and can be broken down in sub-phases. First, around 200 ms after the onset of a word, its category is identified (i.e., whether it is likely to be a verb, a noun, an article, etc.). Subsequently, information of its lexical meaning is retrieved from semantic memory implemented in the middle temporal gyrus. This process takes place approximately between 150 and 300 ms after the onset of a word.

Phase 2 corresponds to a time frame between 300 and 500 ms after stimulus onset. It roughly corresponds to integration, as introduced in the previous section. In this phase, the lexical representation of a word retrieved in Phase 1 is embedded in the contextual representation consisting of the retrieved and unified representation of the other words composing the sentence that is processed (Berkum et al., 1999; Hagoort & van Berkum, 2007; Hagoort et al., 2009; Kutas & Federmeier, 2011).

6.2 Computational models

For the purpose of this study, I use two broad families of computational models developed for word representation generation: non-contextualized models and contextualized models. All these models create so-called word embeddings which consist of vectors of real numbers populating a high-dimensional space. In other words, a model M takes a word w and returns a real vector $\vec{w}$ representing w in an high-dimensional semantic space SS.

The first type of model generates representations $\vec{w}$ that are independent from the context (sentence, paragraph etc.) in which the represented word w is located. I call these types of models **non-contextualized**, and they are represented by the popular word2vec model, see Section 6.2.1 and Mikolov et al., 2013.

(1) In order to open a new account, you should go to a ***bank***.

(2) A fisherman is sitting with his rod on the ***bank*** of the river Thames.

(3) The domestic ***dog*** is a member of the genus Canis, which forms part of the wolf-like canids.

(4) I took my ***dog*** out for a walk in the park.

Besides non-contextualized models, we also consider a contextualized model: ELMo (Section 6.2.2, Peters et al., 2018). This model, contrary to word2vec, assigns representations $\vec{w}$ that depend on the textual context in which the represented word w is located. So, for instance, if the word *dog* appearing in Sentences 3 and 4 always obtains the same embedding $\vec{dog}^{w2v}$ from word2vec, it will obtain two different vectors ($\vec{dog}_3^{ELMo}$ and $\vec{dog}_{@s4}^{ELMo}$), one for each of the two contexts in which it is found.

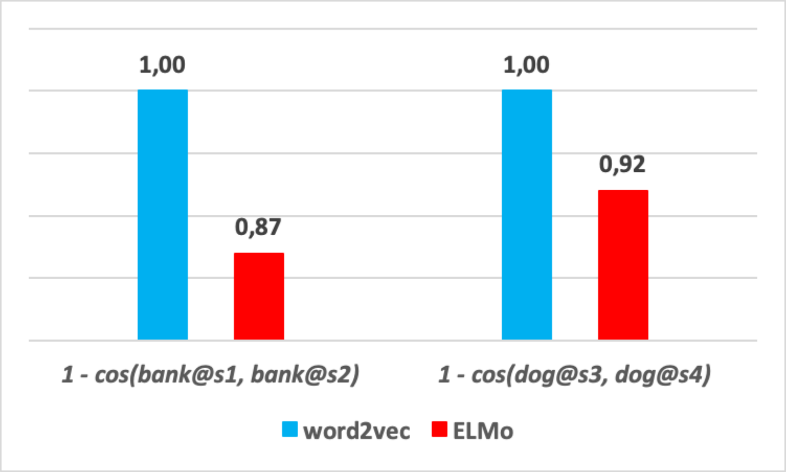


Figure 6.1: Word2vec (blue) return identical representations for words independently from the sentences they are presented in. This is evident from the 1.0 cosine similarity between *bank* in Sentences 1 and 2 – on the one hand – and between *dog* in Sentences 3 and 4. Conversely ELMo (red) instead produces different contextualized representations of the same words depending on the context they are used, note the lower similarities (0.92 and 0.87).

As shown in Figure 6.1, this becomes evident when we compute the similarity between the embeddings. The cosine similarity between the word2vec generated $\vec{dog}_{@s3}^{w2v}$ vector (Sentence 3) and $\vec{dog}_{@s4}^{w2v}$ (Sentence 4) is 1.0, indicating a total identity between the two representations. Instead the cosine similarity between the ELMo generated $\vec{dog}_{@s3}^{ELMo}$ vector (Sentence 3) and $\vec{dog}_{@s4}^{ELMo}$ (Sentence 4) is 0.92. Similarly the similarity between the word2vec representations of bank (Sentences 1 and 2) is 1.0, whereas the similarity between the ELMo representation of the same words is 0.87.

6.2.1 Non-contextualized embeddings (Word2vec)

Word2vec is an artificial neural network-based model used to produce word embeddings. It has been proposed as a more advanced alternative to earlier distributional semantic vector spaces such as latent semantic analysis. From an architectural point of view, word2vec consists of a shallow, two-layer neural network. The model can be trained either to predict the current word from a window of surrounding context words (continuous bag-of-words, CBOW), or –

conversely – predict the surrounding window of context words given a target word (continuous skip-gram, CSG).

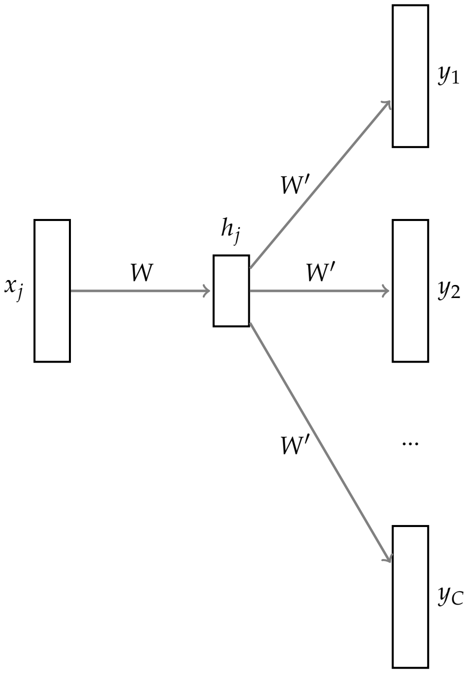


Figure 6.2: Word2vec skip-gram architecture

Training creates a high-dimensional vector space populated by word vectors, which are positioned in the space in such a way that words that share similar semantic and syntactic properties lay in close proximity to one another. Figures 6.2 and 6.3 exemplify the architectural alternatives for word2vec training (CBOW and CSG) (adapted from Hassan & Mahmood, 2017). Note the symmetrical structure of the two architectures. In the case of the CBOW variant, which is used in the present chapter, inputs x_1, x_2, and x_C refers to the 1^{st}, 2^{nd} and C^{th} words in the window surrounding target word y_j that the model has to predict during training, and h_j is the hidden layer. In both cases the trained embeddings correspond to the weights stored in matrix W, whose dimensions $[v \times d]$ correspond to the size of the modelled vocabulary (v) and the chosen number of dimensions (d) of the vector space itself. Once the model is trained, the way word2vec assigns embeddings to word w can be seen as a sort of dictionary "look-up" where the embedding of word w_i corresponds to row i of matrix W.

For the purpose of the study presented in this chapter, it is important to point out the role of context with regard to the way word2vec is trained and used to assign word embeddings.

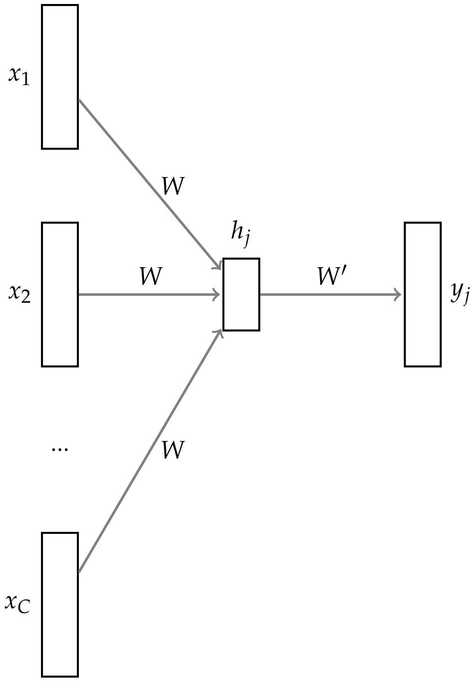


Figure 6.3: Word2vec bag-of-word architecture

Context indeed plays a capital role during the training of the model, both in the case of the CBOW and of the CFG variants. In both cases the context of a word is present in the pipeline, either as the input or as the target of the training function. Nonetheless, once the model is trained, its application is blind to the context and relations that the words have.

6.2.2 Contextualized word embeddings (ELMo)

The contextualized word embedding model ELMo (Peters et al., 2018) relies on the properties of recurrent neural networks. Contrary to word2vec, ELMo is a deep contextualized model for the generation of word representation. It models complex characteristics of word use that vary across linguistic contexts.

From an architectural point of view, ELMo is a recurrent bi-directional language model (biLM) composed of 2 layers of bi-directional recurrent units (implemented as LSTM layers) feeding to an output softmax layer. A language model refers to a system (e.g., stochastic or neural) trained on predicting a word given its preceding context. In its most common formulation, a language model is an approximation of the function $P(w_i|w_{1:i-1})$, i.e. a function that assigned a probability to a word w_i given its prefix $w_{1:i-1}$.

A recurrent layer is a layer that creates a representation of the input sequence (sentence) at word w_i as a combination of the representation of the i_{th} word and the representation of its preceding (if the recurrent neural network proceeds left to right) or following (if it proceeds right to left) context.

Figure 6.4 illustrates the structure of the ELMo model consisting of a 2-layer biLM. ELMo's bi-directionality refers to the fact that its hidden layers receive information about both the preceding and following words of w_i. The representation ELMo generates for a word consists of a concatenation of the activations of its two recurrent layers.

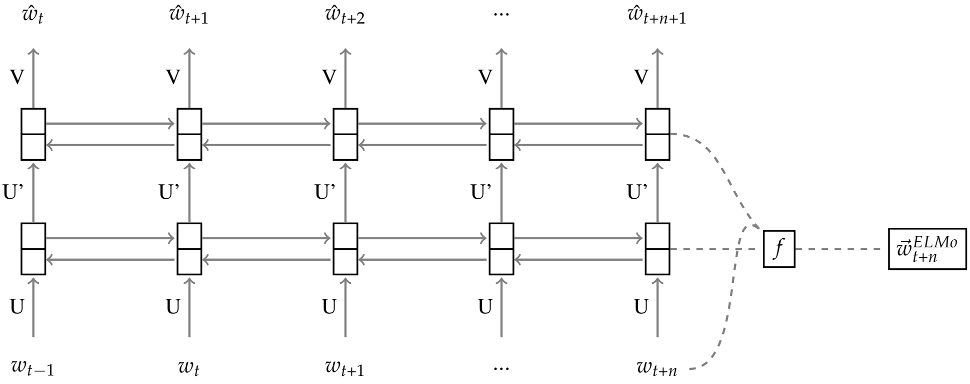


Figure 6.4: ELMo architecture consisting of a 2-layer biLM, where the layers are implemented as LSTMs. The contextualized embedding is produced as a linear combination of its components.

Contrary to how word2vec assigns embeddings to words, given word w_i in a sentence $S = [w_1, w_2, ...w_n]$, ELMo instead creates the embedding $\vec{w}_i$ by passing the whole S text through the biLM. Embedding $\vec{w}_i$ will depend on context $w_{1:i-1}$ and $w_{i+1:n}$ as the combination of the activation of the LSTM layers corresponding to the presentation of w_i to the model. For these reasons, if w_i appears in a different sentential context S', its embedding $\vec{w}_i{}'$ will be different.

6.2.3 Training

Since the MEG data (described in Section 6.4) was collected from Dutch native speakers reading Dutch sentences, I used models trained on Dutch texts. Word2Vec was trained on the CBOW task on the Corpus of Spoken Dutch (CGN2). For ELMoI used the pretrained

Dutch model provided by the ELMoForManyLangs collection[1] (Che et al., 2018; Fares et al., 2017) with the same hyperparameter settings as Peters et al., 2017 and trained on a set of 20-million-words Dutch corpus randomly sampled from the raw text released by the CoNLL 2018 shared task on Universal Dependencies Parsing.

6.3 Hypotheses

This study aims at investigating the neurobiological underpinnings of the distinction between non-contextualized and contextual representation as derived from recent developments in word embedding modeling techniques. More specifically, for the non-contextual model, I hypothesize a similarity between this model and the lexical retrieval operation of human language processing. I expect that a model with such characteristics should most closely resemble the activity in lateral temporal regions during processing, around 200 ms after word onset.

Conversely, contextual model representations of words are expected to resemble more the activity related to integration in inferior frontal and antero-temporal regions between 300 and 500 ms after the onset of the word.

6.4 MEG Data

For the purpose of this study, I have used the MEG described in Chapter 2, Section 2.3. The data belongs to the MOUS dataset (Schoffelen et al., 2019) collected at the Donders Centre for Cognitive Neuroimaging in Nijmegen, The Netherlands. For more details on the acquisition procedure, stimuli, preprocessing and source reconstruction techniques, I refer to the original chapter and to Lam et al., 2016.

6.5 Analysis

The computational models introduced in Section 6.2 are used to generate vectorial representations of the same stimulus sentences presented during the acquisition of the magnetoencephalographic data described in Section 6.4. These representations are generated at word level, meaning that these models assign a set of real number vectors for each word of the stimulus sentences. The non-contextualized word embedding (word2vec) assigns only one

---

[1] `https://github.com/HIT-SCIR/ELMoForManyLangs`

vector per word, whereas the contextualized models (ELMo) each represent every word with a set of vectors, each corresponding with one component (layer) of their internal architecture.

In these analyses, a word represented by ELMo is assigned one vectorial representation corresponding to the average of the activation of the two bi-directional layers composing the ELMo network ($\vec{ELMo}$).

This section describes the analysis methodologies employed to map the two vectorial representations ($\vec{W2V}$, $\vec{ELMo}$) to the corresponding brain activity recorded with MEG. The goal is to both map the overlap between model representations and brain activity at the anatomical level, and to track the temporal evolution of such similarity. For this reason, I adopted a version of representational similarity analysis (RSA) implemented in such a way to return both spatially (anatomical) and temporally situated similarity scores.

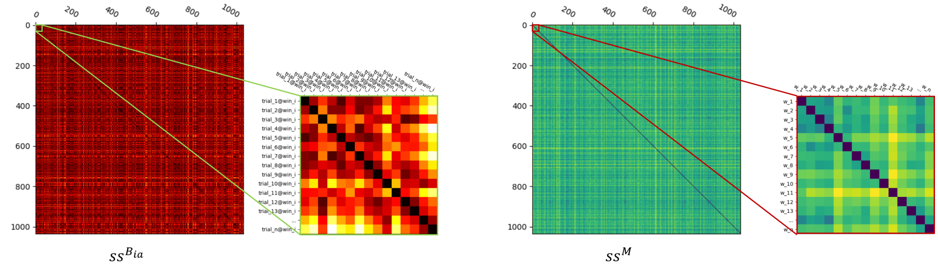
6.5.1 Representational similarity analysis through time

Given a set of linguistic units, for instance, words $w_1, ...w_n$, a vectorial representation for each of them can be generated using the embedding models described in Section 6.2. Words $w_1, ...w_n$ are assigned vectors $\vec{w_1}^M, ...\vec{w_n}^M$ from the computational model M. These are drawn from a representational space RS^M populated by the vectorial embeddings. At the same time, it is possible to derive a brain representation of the same words. This representation is the word-wise trial activation recorded in the MEG dataset described in Section 6.4. Note that the signal for each word was not average over time per trial, but that it was kept as is. These activation samples also create a representational space RS^B populated by the same words $w_1, ...w_n$. At this point, the aim of the analysis is simply to measure how similar RS^M and RS^B are. In order to do so, representational similarity analysis (RSA) was used. Instead of directly mapping the two spaces, RSA compares pairwise similarities across different spaces.

RSA was conducted at the level of anatomical regions and using a sliding, partially overlapping temporal window (width 30 ms, step 16 ms). In this way, a neural representation is obtained for each word for each brain region and it is paired with a specific model-derived vectorial representation of the same word. This allows to plot the similarity score between model and windowed time-course of each anatomical region of the MEG dataset.

Given a model RS^M, a representational similarity structure ss^M consists of a $[n \times n]$ matrix where element ss^M_{ij} denotes the dissimilarity (computed as Euclidean distance) between embeddings $\vec{w_i}^M$ and $\vec{w_j}^M$ from RS^M of words w_i and w_j. Similarly it is possible to derive a series of representational structures $ss^{B_{rt}}$ for each anatomical region r and time window

t. Element $ss_{ij}^{B_{rt}}$ quantifies the similarity between the brain activity in r and time t for word stimuli w_i and w_j. Figure 6.5 displays the representational similarity structure for brain activity in region r and time i 6.5a, and for model M 6.5b.



a **b**

Figure 6.5: Similarity matrix for brain activity in region a and time i (a), and for model M (b). The matrices have been zoomed in to show their fine-grained structure and the labels of their rows and columns.

The similarity score is estimated by taking Pearson's correlation coefficient between the upper off-diagonal triangle of the $[n \times n]$ symmetric paired similarity matrices (ss^M and $ss^{B_{ai}}$) (Figure 6.6). These scores quantify the extent to which the similarity across stimuli is similarly represented by the model M and by brain activity in anatomical region a and time i. These measures are repeated across time t and anatomical regions r (Figure 6.7).

Therefore for each anatomical region, I obtained a representation of similarity between model and brain activity as a function of time from 0 to 500 ms after word onset (in windows of 30 ms, every 16 ms).

The analyses are conducted at the single-subject level up to this stage. The result of the temporal RSA for model M thus consists of 74 matrices, one for each subject, of size $[s \times u]$ where s is the number of anatomical regions and u the number of temporal windows.

6.5.2 Group-level analysis

Given 74 per-subject $[s \times u]$ matrices, I want to obtain a single group-level matrix describing the similarity between model M and brain activity across anatomical regions and over time.

Typically, group-level RSA results are obtained by averaging similarity matrices across subjects before computing the similarity score with the model similarity matrix. Cross-subject averaging requires that all matrices have the same size and that the row and columns correspond to the same words across matrices. This is not the case for this dataset. First, the subjects are

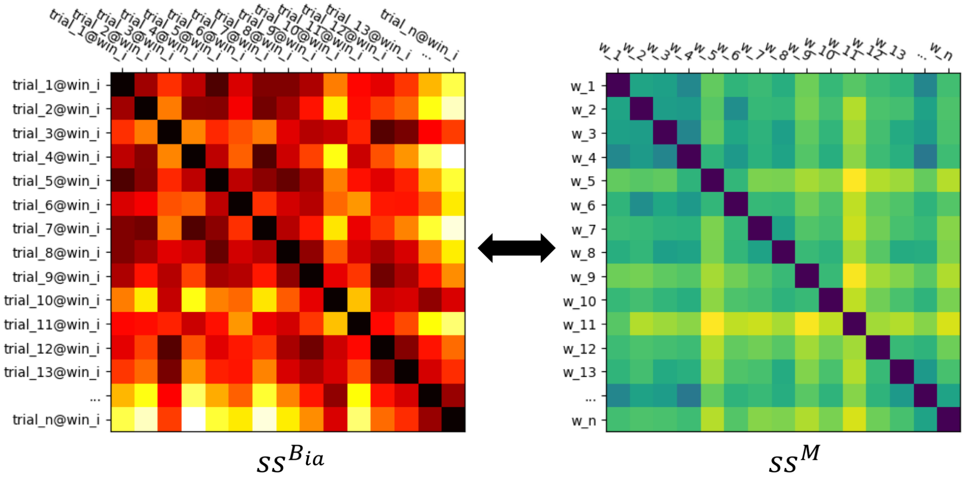


Figure 6.6: RSA is conducted by computing the similarity score between the brain space similarity matrices (left) and model similarity matrix (right). Here, the brain similarity matrix consists of the pairwise similarity scores among trials at anatomical region a and time window i.

grouped in cohorts of about 20, each of which was presented with a different subset of the 360 stimulus sentences. This means that not all $ss^{B_{rt}}$ have rows and columns corresponding to the same words. Moreover, the trial selection procedure in the per-subject pre-processing consisted of discarding trials with irreparable artifacts. Therefore, even if two subjects are presented with the same set of sentences, there is a chance that they would have a different set of corresponding MEG trials due to the different occurrence of signal artifacts.

In this analysis, I computed t-statistics over subjects for each region/time combination independently. I computed a one-sided t-score for each of the present computational models ($\vec{W2V}$, $\vec{ELMo}$). I also computed a one-side t-score between the scores obtained by the aggregate models and word2vec ($\vec{ELMo} > \vec{W2V}$). The results of these t-statistics are thresholded at $p < 0.05$. Given the exploratory nature of this study, no statistic correction was applied in order not to obliterate the possible small effects detected by the RSA.

6.6 Results

The results of the analyses are split in two main parts. In the first part, I report the results of each of the two embedding models (word2vec and ELMo) separately. In the second part of the analyses, the similarity scores of word2vec and ELMo are contrasted.

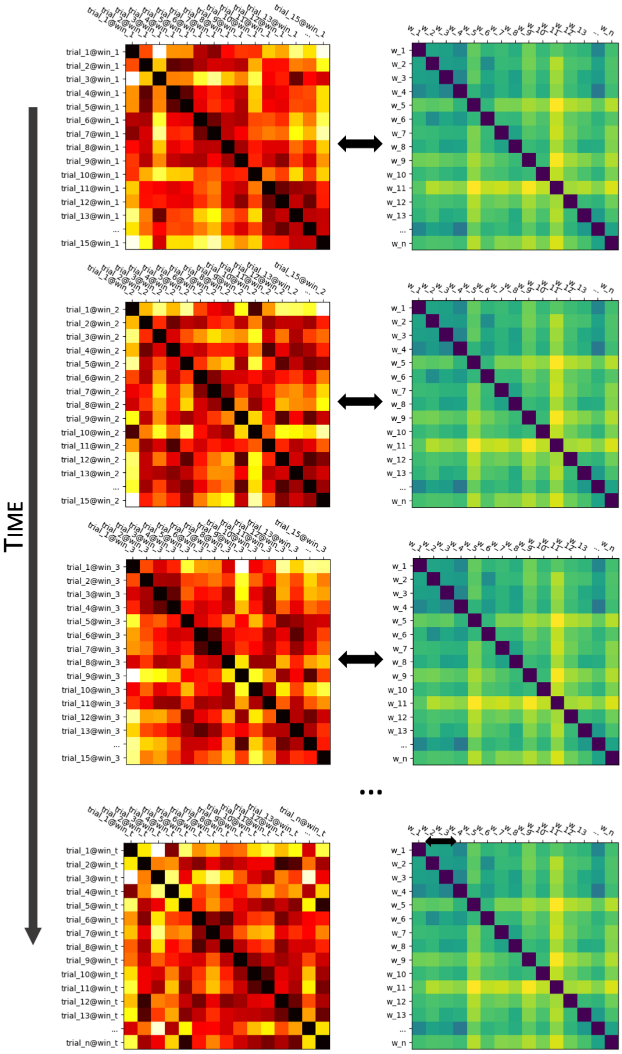


Figure 6.7: RSA is conducted over time, meaning that a brain space similarity matrix is computed for each separate time windows slicing the MEG trials.

Results are provided at the whole-brain level, displaying the model–brain similarities at 5 distinct time points: 150, 250, 350, 450, and 550 milliseconds after word onset.

6.6.1 Models

The RSA analyses of the two computational models are reported separately in Figure 6.8 and 6.9.

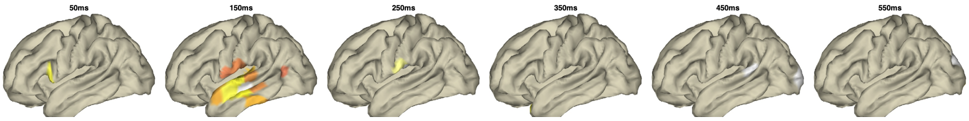


Figure 6.8: Whole brain similarity over 6 time frames for word2vec.

The embedding model (Figure 6.8), which does not include contextual information, word2vec, returns lower similarities overall over time when correlated with brain activity, but more so from about 300 ms post word onset. For earlier latencies, word2vec shows significant activity in the left middle and inferior temporal gyri. Significant similarity with brain activity is also observed around 400 ms in the left posterior superior temporal gyrus.

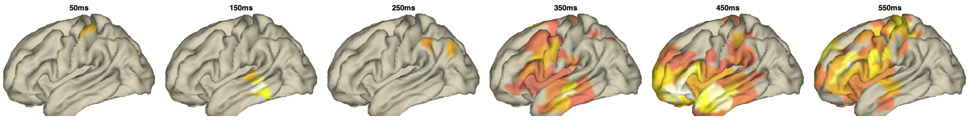


Figure 6.9: Whole brain similarity over 6 time frames for ELMo.

The contextualized model (Figure 6.9), ELMo, instead exhibits an overall significant similarity with brain activity between 300 and 500 ms in the left frontal, prefrontal, and anterior temporal regions. In particular, ELMo shows significant similarity with the left inferior temporal gyrus and the left anterior temporal cortex around 400 ms.

6.6.2 Comparison

In Figure 6.10, I show the results of a direct comparison between ELMo and word2vec ($\vec{ELMo} > \vec{W2V}$).

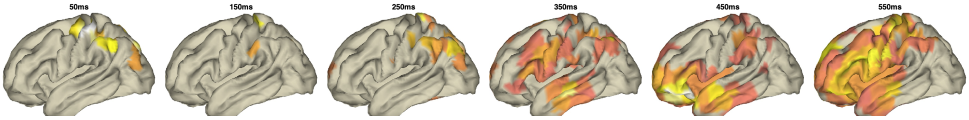


Figure 6.10: Comparison between contextualized model (ELMo) with word2vec (ELMo > word2vec)

ELMo shows significantly higher similarity to brain activity as compared to word2vec in the left anterior temporal lobe and in the left inferior frontal gyrus around 400 ms post word onset. The higher scores are also observed for the left middle frontal and prefrontal regions around 500 ms.

6.7 Discussion

In Section 6.6, I observed that contextual and non-contextualized models yield qualitatively different results with regard to the timing and location of their similarity to MEG-recorded brain activity. In this section, I discuss the implications of these findings in light of the nature of the models and of the brain processing of natural language, as introduced in Sections 6.1.1 and 6.1.2.

I believe that computational word embedding models help in probing the nature of the neural representations correlated to memory retrieval and to integration. This is because they make the distinction between these two phenomena more computationally specific. When discussing the nature of retrieval and integration, Baggio & Hagoort, 2011 argue that approaches based on formal semantics might not be realistic models of how the brain implements. In agreement with Seuren, 2009, they state that formal semantics disregards natural language as a psychological phenomenon. They continue stating their desire to develop an account "that adheres to cognitive realism, in that it explains how language users derive meaning and how the human brain instantiates the neural architecture necessary to achieve this feat". I believe that distributional semantic models, of which contextualized embeddings are the most advanced version, have already been proven their cognitive realism by being good models of human behavioral – e.g. semantic similarity judgment (Baroni & Lenci, 2011) – and neural

(Mitchell et al., 2008; Anderson et al., 2015) data. Moreover, at the dawn of the field, distributional models – e.g., Latent Semantic Analysis (Landauer & Dumais, 1997) – were actually developed as cognitive models to answer questions on how infants acquire word meaning, and humans react to semantic similarity and relatedness. In light of the above considerations, I think that the models presented in the study might offer a cognitively realistic approximation of what is going on in the brain during memory retrieval and integration.

In the remainder of this section, I will discuss the effect of contextualization on the similarities between computational representations and brain activity (Section 6.7.1). I will specifically focus on the implications of these findings regarding the role of the anterior temporal lobe 6.7.2 and of activity peaking around 400 ms after stimulus onset(Section 6.7.3). I will also discuss the plausibility of the models chosen for the present study (Section 6.7.4).

6.7.1 The effect of contextualization on model-brain similarity

ELMo creates vectorial representation of words (embeddings) dependent on the context in which those words occur (Section 6.2.2). This contrasts with the nature of embeddings generated by a model such as word2vec (Section 6.2.1) that, once trained, always generates the same embeddings for a word regardless of its context.

Contextualized word embeddings generated by ELMo display higher similarity with brain activity in the frontal and anterior temporal regions of the brain around 300 and 400 ms post word onset; in other words, in areas and in a time frame compatible with integration processes. Integration refers to the integration of a linguistic unit (a word for instance) in the context provided by the other linguistic input in which it happens to be contained.

Non-contextualized word embeddings generated by word2vec exhibit a somewhat opposite behavior, showing higher similarity with earlier (100-200 ms) activity in lateral temporal regions. These regions are supposed to implement long-term memory (semantic memory) retrieval.

The fact that ELMo and word2vec embeddings display such different behaviors with regard to brain activity can be reconciled with the division of labor predicated by models such as the MUC (Hagoort, 2013a), especially for what concerns the distinction between semantic memory and contextual integration.

6.7.2 The role of the anterior temporal lobe in integration

The fact that contextualized embeddings show higher similarity in the left anterior temporal lobe might indicate that this region plays a role in integration. This is in line with several experimental studies reporting the involvement of this area in lexical, semantic, and syntactic integration. Works such as Mazoyer et al., 1993, Stowe et al., 1998, Friederici et al., 2000, Humphries et al., 2006, and Humphries et al., 2007 arrived at the conclusion that the anterior temporal lobe is somehow involved in sentence-level integration, after observing an increase in activity in this area during the presentation of sentences as compared to word lists. In addition, a series of other studies have confirmed the role of ATP in processing composition by showing that its activity is modulated by the type of syntactic relation holding between two words being integrated (Baron et al., 2010; Baron & Osherson, 2011; Bemis & Pylkkänen, 2011, 2013; Westerlund et al., 2015; Bemis & Pylkkänen, 2013).

The left ATP is also considered central in semantic memory. This was confirmed by studies on patients suffering from semantic dementia (Hodges et al., 1992, 1995; Mummery et al., 2000; Rogers et al., 2004), and several functional neuroimaging studies (Gauthier et al., 1997; Tyler et al., 2004; Bright et al., 2005; Moss et al., 2004; Rogers et al., 2006). These findings are summarised by Patterson et al., 2007, which hypothesizes that concepts are represented by a network of sensorimotor representations converging in the left ATP, which integrates modality-specific features in order to produce supra-modal representations.

Contextualized models do not show high correlations in the ATP at latencies that are associated with memory-related processing: 100-200 ms after word presentation. The results, therefore, seem to indicate that contextualized embeddings approximate representations that have to do with the integration into context and not with lexical retrieval from memory per se. If this were the case, I would have expected similar correlations in the ATP both at memory retrieval (100-200 ms) and integration latencies (400 ms). Conversely, a tentative interpretation with regard to the role of ATP could be that activity in this region is different when semantic memory operations and contextual integration processes are carried out.

6.7.3 The role of N400 in integration

The results presented in Section 6.6 do confirm not only the anatomical loci of memory and integration, but also provide indirect suggestions on the role of the N400 (Kutas & Federmeier, 2011). It has been debated in the literature whether the N400 is best characterized as playing a role in combinatorial (integration) or non-combinatorial (retrieval from memory) processes. Baggio & Hagoort, 2011 and Hagoort et al., 2009 refrain from providing a stance on the matter because of the difficulty of devising convincing task-dependent experimental designs that are able to disentangle combinatorial and non-combinatorial semantic processes. Here again, I believe that the present computational modeling approach may provide an answer to the question. The results seem to point more towards a combinatorial process for the N400, given that the contextualized model, which represents the result of contextual integration to predict the word in focus, maps to brain activity corresponding to its latencies. It is an indirect proof, but a proof pointing in this direction nonetheless. These results are in line with studies that link the role of the N400 to context processing (Berkum et al., 1999; Hagoort & van Berkum, 2007).

6.7.4 The plausibility of bi-directional RNNs

ELMo is essentially a bi-directional recurrent neural network-based language model that integrates a word with its preceding and following context. Bi-directional recurrent language models seem to violate the assumption that human language processing proceeds left-to-right and word-by-word. Although this is trivially true for listening, it is worth noting that studies of reading behavior tend to describe a more nuanced situation. It is consistently reported in the eye-tracking of reading literature that between 15 and 20 % of eye movements proceed right-to-left (eye regressions) for left-to-right languages such as English or Dutch, and the opposite occurs for right-to-left languages such as Arabic (Rayner & Pollatsek, 1995). Moreover, a number of "jump-ahead" eye movements are also commonly observed, indicating that humans either skip information that is deemed irrelevant for the processing of a linguistic item or that they look ahead in order to collect contextual information to the right of a word. This indicates that the preceding linguistic sequence is not always the only contextual material employed in the processing and interpretation of a word or of a sentence as a whole.

6.8 Conclusions

Recent developments in computational linguistics have created a new family of models that generate word embeddings that compute their representations with information derived from the context in which words are used. In this chapter, I have adopted one particular contextualized model, ELMo, as an approximation of the result of integration processes in the human brain during natural language comprehension. I contrasted ELMo with word2vec, a non-contextual embeddings model. Starting from the distinction between semantic memory retrieval (implemented in temporal regions and activated around 200 ms after the onset of an incoming word) and word integration into context (carried out in inferior frontal, perisylvian and anterior temporal regions around 400 ms after word onset), I observed that non-contextualized models correlate with activity only in regions and latencies associated with semantic memory. In contrast, contextualized models correlate with activity in areas and latencies associated with word integration in context. These results confirm the functional and physiological distinction between memory and integration. Moreover, they provide some insight into the role of the IFG, an area involved in integration and whose activity might temporarily store contextualized lexical representations. The results also point towards an involvement of the anterior temporal lobe in integration, an area that was already linked to semantic combinatorial processes and which nonetheless received less attention in the theories of the architecture of the language system adopted in this chapter.

By highlighting a parallelism between models and brain activity, my results offer a contribution to the understanding of the division of labor at the cortical level between areas encoding lexical items in isolation and areas sensitive to the use of those items in context.

7

Conclusions

I began this dissertation with the assumption that language processing in the brain can be partitioned into several sub-components that can be modelled by different computational linguistics models fitted on human data collected using a naturalistic stimulus paradigm. In this final chapter, I summarize the main results with regards to the sequential properties of the stimulus, syntactic analysis, and basic sub-operations, I also provide some insights derived from the application of computational tools and brain imaging data collected using naturalistic stimuli.

In this dissertation, I set out to investigate the distinct sub-processes constituting language processing in the brain using naturalistic language stimuli and computational modeling. In this chapter, I summarize the results exposed in the previous chapters, and I draw conclusions with regard to the methodological choices that were adopted throughout the whole research presented herein.

I investigated whether different sequential properties of language, types syntactic structure, and sub-operations can be distinguished using computational modeling applied to naturalistic stimuli.

The first question dealt with the **sequential properties of the stimulus**. In Chapter 3, I addressed the questions of whether the brain is sensitive to the sequential properties of the stimuli defined in terms of stochastic language models, and whether the sequences of phonemes, words and grammatical categories constituting the stimulus are processed by distinct areas within the brain language network.

Subsequently, in Chapters 4 and 5, I investigate the type of syntactic formalisms that might describe the **structural analyses** conducted during language processing. In particular, in Chapter 4, I investigated the type of grammars that better describe the activity in a set of cortical areas hypothesised to be involved in syntactic analyses. In Chapter 5, instead, my results showed that it is possible to find evidences for syntactic parsing by looking at patterns of eye movements during reading.

Finally, concerning the **basic operations** supporting language processing and in particular the role of contextualization, in Chapter 6, I addressed the question of whether lexical retrieval and integration can be formalized in terms of contextualization of linguistic representations (i.e., word embeddings). I used contextualized word embeddings as models of integration processing, and conversely, non-contextualized word embeddings as models of lexical retrieval in the brain.

7.1 Overview of the results

Before discussing the theoretical and methodological implications of the studies described in the previous chapters, the present section offers an overview of the results contained in this dissertation. The central motif of this work is the mapping between computational linguistic model-derived measures and brain activity in several cortical areas. Therefore, in this section, I present two tables: Table 7.1 lists for each brain area all the computational models that map or explain its activity; Table 7.2 contains the same information but, instead, lists for each model its corresponding brain areas.

Table 7.1: Area-wise summary: list of brain areas mentioned in this dissertation and their correspondent computational models.

Area	Model(s)	Function	Section(s)
left angular gyrus	SLM-phoneme	phonemic sequence	3.3.3
left anterior temporal pole	SLM-word	lexical sequence	3.3.1
	DG	shallow structure	4.6.2
	Contextualized	lexical integration	6.5.1
left fusiform gyrus	SLM-word	lexical sequence	3.3.1
left inferior frontal gyrus	DG	shallow structure	4.6.2
	Contextualized	lexical integration	6.5.1
left inferior parietal lobule	SLM-phoneme	phonemic sequence	3.3.3
left inferior temporal gyrus	SLM-word	lexical sequence	3.3.1
left insula	SLM-phoneme	phonemic sequence	3.3.3
left middle superior frontal gyrus	SLM-PoS	grammatical sequence	3.3.2
left middle temporal gyrus	SLM-PoS	grammatical sequence	3.3.2
	SLM-phoneme	phonemic sequence	3.3.3
	Non-contextualized	lexical retrieval	6.5.1
left posterior superior temporal gyrus	SLM-word	lexical sequence	3.3.1
	PSG	phrase structure	4.6.2
left precentral sulcus	SLM-PoS	grammatical sequence	3.3.2
right amygdala	SLM-PoS	grammatical sequence	3.3.2
right angular gyrus	SLM-PoS	grammatical sequence	3.3.2
	SLM-phoneme	phonemic sequence	3.3.3
right Heschl's gyrus	SLM-phoneme	phonemic sequence	3.3.3
right middle temporal gyrus	SLM-PoS	grammatical sequence	3.3.2

right posterior middle temporal gyrus	SLM-phoneme	phonemic sequence	3.3.3
	SLM-word	lexical sequence	3.3.1
right precentral sulcus	SLM-PoS	grammatical sequence	3.3.2
right putamen	SLM-PoS	grammatical sequence	3.3.2
right superior frontal gyrus	SLM-phoneme	phonemic sequence	3.3.3

Table 7.2: **Model-wise summary: list of computational models adopted in this dissertation and their correspondent brain areas.**

Type	**Model(s)**	**Area**	**Section(s)**
Word Embeddings	Contextualized	left anterior temporal pole	6.5.1
		left inferior frontal gyrus	6.5.1
	Non-contextualized	left middle temporal gyrus	6.5.1
Stochastic Language Model	SLM-phoneme	left angular gyrus	3.3.3
		left inferior parietal lobule	3.3.3
		left insula	3.3.3
		left middle temporal gyrus	3.3.3
		right angular gyrus	3.3.3
		right Heschl's gyrus	3.3.3
		right posterior middle temporal gyrus	3.3.3
		right superior frontal gyrus	3.3.3
	SLM-PoS	left middle superior frontal gyrus	3.3.2
		left middle temporal gyrus	3.3.2
		left precentral sulcus	3.3.2
		right amygdala	3.3.2
		right angular gyrus	3.3.2
		right middle temporal gyrus	3.3.2
		right precentral sulcus	3.3.2
		right putamen	3.3.2
	SLM-word	left anterior temporal pole	3.3.1
		left fusiform gyrus	3.3.1
		left inferior temporal gyrus	3.3.1

		left stuperior temporal gyrus	3.3.1
		right posterior middle temporal gyrus	3.3.1
Syntactic Parser	DG	left anterior temporal pole	4.6.2
		left inferior frontal gyrus	4.6.2
	PSG	left stuperior temporal gyrus	4.6.2

7.1.1 Sequential processing of language at different information levels

In Chapter 3, I demonstrated that areas sensitive to the sequential properties of language could be partitioned in sub-networks, each of which is sensitive to different types of information. These analyses showed that the stochastic sequential processing paradigm is indeed a powerful formalism able to predict neurobiological correlates in areas belonging to the language processing network, also when applied to sub-lexical (phonemic) and syntactic (part of speech) levels.

Three distinct sets of cortical areas are sensitive to the three types of sequential information – lexical, syntactic, and phonological – examined in this dissertation. This confirmed the hypothesis that language processing could indeed be decomposed into different levels of information corresponding to different subdivisions of the language network. No area was activated by all three types of information, and only limited sets of voxels show overlap between pairs of types of information. Activity in the inferior and lateral portions of the left temporal cortex is likely involved by **lexical** sequential information and is likely to be a central hub of lexical information processing (Binder et al., 2009; Binder & Desai, 2011). **Phonological** information instead concerns activity in the temporal cortex only in regions close to the transverse gyrus (Heschl's gyrus) (Mendoza, 2011), and the supplementary motor area (SMA) (Alario et al., 2006; Willems & Hagoort, 2007; Hertrich et al., 2016). The **syntactic** sequential information explains activity centered in the middle temporal gyrus (MTG).

In conclusion, these results appear to confirm the intuition that language is processed in parallel by distinct networks sensitive to different sources of information, including at least those tested here: phonological, lexical, and syntactic.

7.1.2 Syntactic analysis: dependency and phrase-structure processing in the brain

In Chapter 4, I investigated whether different brain regions are sensitive to different kinds of syntactic structures. In order to do so, I assessed dependency (DG) and phrase-structure (PSG) descriptors of sentences as predictors of brain activity in the left anterior temporal pole (ATP), the left inferior frontal gyrus (IFG), and the left superior temporal gyrus (STG).

The results indicate a general division of labour between areas sensitive to PSG in the posterior perisylvisn cortex (left pSTG), and areas sensitive to DG measures in more anterior cortical regions (left IFG and ATP). As I will discuss in Section 7.2, the fact that activity in the left ATP is better explained by DG measures as compared to PSG ones is in accordance with Westerlund et al., 2015, and corroborates the idea that ATP works as a hub for sentential-level semantic composition where words are combined according to the argument structure of the sentences as captured by their dependency parses.

7.1.3 Dependency structures and eye movements

Syntactic analysis was also investigated using data other than brain imaging. In Chapter 5, I have presented a study investigating whether eye movements of readers may be influenced directly by the syntactic structure of the sentence. I tested this using structures computed by a dependency parser. The hypothesis was that the path of regressions from a word to an earlier word coincide, at least partially, with the edges of dependency relations between these words. The results indicate that there is a significant effect of the number of left-hand side dependency relations on the number of backward saccades. These results are corroborated by the observation that about 40% of backward saccades do indeed land on target words engaged in dependency relations according to the syntactic structure of the sentences composing the stimuli.

Altogether these results converge on the idea that eye movements reflect, among other things, the shallow syntactic structure of language. Moreover these results also seem to corroborate the idea that humans do engage in online syntactic analysis of the input – at least in the form of dependency parsing.

7.1.4 Tracing operations underpinning language processing

In Chapter 6, I investigated the difference between lexical retrieval and integration using contextualized (e.g., ELMo – Peters et al., 2018) and non-contextualized (e.g., word2vec – Mikolov et al., 2013) word embeddings, models that create numerical representations of linguistic units on the basis of their use in large corpora of texts. Contextualized embeddings create representation that depends on the actual context a specific token is uttered and are used in this study as an approximation of the result of integration processes in the human brain during natural language comprehension. On the other hand, non-contextualized models represent tokens independently of the context of utterance and are used as approximation of lexical retrieval.

The results show a correspondence between non-contextualized embeddings and activity related to lexical retrieval, and between contextualized embeddings and activity associated instead with lexical integration. In more details, non-contextualized embeddings correlate with activity around 200 ms post word onset in the temporal cortex, whereas contextualized embeddings correlate mostly with activity after 400 ms post word onset in anterior temporal and frontal regions. These results confirm the functional and physiological distinction between memory and integration. Moreover, they provide some insight into the role of the IFG and the involvement of the ATP in integration. This latter area that was already linked to semantic combinatorial processes and which nonetheless received less attention in the theories of the architecture of the language system that have been considered in this dissertation.

7.2 Remarks on the left anterior temporal lobe

The left anterior temporal lobe appears to be involved in at least two of the three sub-components of language processing in the brain probed in this dissertation: syntactic analysis and lexical integration in context.

The results presented in Chapter 4 seem to indicate that the ATP (together with parts of the IFG) plays a role in the construction of sentence-level representations following a dependency grammar. In fact, in that chapter, we saw how the number of left-hand side dependency relations for each word has a significant effect in the prediction of activity in the left ATP and part of the left IFG (Section 4.6.1). These results are in line with studies such as Mazoyer et al., 1993, Stowe et al., 1998, Friederici et al., 2000, Humphries et al., 2006, and Humphries et al., 2007. Moreover, it is worth noticing how dependency relations are comparable to the 2-word stimuli that were found eliciting activation in the ATP by Baron et al., 2010 and by Westerlund et al., 2015 (adjective–noun, subject–predicate, object–predicate, etc.). In fact, dependency relations making up the structure of a dependency parse can be broadly classified as predicate–argument (e.g., subject–predicate and object–predicate) or modifier (e.g., adjective–noun, adverb–verb, etc.) relations.

In lights of these observations, it seems natural to conclude that the role played by the left ATP in sentence processing is related to the computation of binary relations as captured by a dependency grammar.

Chapter 6 demonstrated that the activity in the left ATP is correlated to the representations obtained from contextualized word embeddings. This was interpreted as an indication that this region plays a role also in sentence-level lexical integration into context.

In conclusion, both the results of Chapter 4 and of Chapter 6 converge on the idea that the left ATP might be the locus where sentence-level semantic representations are computed. These larger representations might be produced by combining the lexical representations of the words composing the sentence following the structure specified in a fashion comparable to a dependency parse, or by integrating lexical information with the information provided by the context of utterance of the stimulus.

7.3 Remarks on naturalistic stimuli and computational modeling

In this dissertation, I have investigated the multiple levels that compose language processing from a sequential point of view, as well as from the point of view of the syntactic structures and of the types of operations involved. In order to do so, I have decided to adopt two innovative and complementary methodological paradigms: naturalistic stimuli and computational linguistic modeling.

The naturalistic stimulus paradigm consists of the collection of data from participants that were presented with stimuli and conditions that resemble as much as possible real-life situations, such as listening to or reading narratives, and with little or no experimental task forced onto them. This is in stark contrast with other approaches that rely on carefully constructed conditions and stimuli. Computational modeling – at least in the way it is meant here – refers to a series of tools adopted from computational linguistics and natural language processing (NLP), which are used to provide a detailed quantitative description of the stimulus.

Data collected using naturalistic stimuli can be a goldmine of information that can be probed by modelling the stimulus. In Chapter 2, I described three distinct datasets, two of which were collected using naturalistic stimuli.

Besides the consideration already expressed in Chapter 1 regarding the advantages of a naturalistic paradigm over more traditional task-oriented approaches, at the end of this dissertation, it would be in order to discuss one of its most recurrent criticisms: the lack of control compared to task-oriented paradigms. As I mentioned in Chapter 1, naturalistic paradigms are expressly distinct from task-oriented paradigms, which are deliberately controlled and decontextualized. The use of naturalistic stimuli does not necessarily entail a complete loss of the experimental control. The potential lack of control from the point of view of the stimulus is instead compensated by modeling the stimulus using computational linguistic tools. This approach leverages computational models in order to account for as much variance as possible in the data by providing descriptors for as many factors concerning the stimulus as possible (i.e. word frequency, position, length, and predictors of interests such as surprisal and syntactic structure).

The naturalistic stimuli datasets presented in Chapter 2 have been used for several different studies ranging from engagement during narrative comprehension, mental imagery, and computations underlying processing (Willems et al., 2016; Frank & Willems, 2017; Lopopolo et al., 2017; Mak & Willems, 2018), showing that the richness of the material presented to the subjects allows for the study of several – even quite different – aspects of language processing with a single dataset. Therefore, another benefit is the fact that this approach encourages the *re-use* of data, which is expensive and time-consuming to collect.

7.4 Final remarks

In this dissertation, I have investigated the hypothesis that language processing in the brain can be partitioned into several sub-components using different computational linguistics models and human data collected under a naturalistic stimulus paradigm.

Different types of sequential information derived from the stimuli can be traced in distinct cortical areas belonging to the language network in the brain. At the same time, different types of syntactic operations seem to be performed by distinct areas in the brain sensitive to syntax. Finally, contextualization of lexical information can be modelled using state of the art computational word embeddings, which can help to understand the distinction between lexical retrieval and integration, both from an anatomical and temporal point of view.

This dissertation offers a contribution to the potential derived from combining unconstrained naturalistic stimulus paradigms with the modeling of the stimulus derived from computational linguistics tools. The full potential of this approach remains to be fully tested, since the rapid evolution and innovation in computational linguistics offers new models every day that could potentially be used for further testing hypotheses about human language processing.

Bibliography

Alario, F.-X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (sma) in word production. *Brain Research*, *Mar 3*(1076(1)), 129-143.

Altmann, G. T., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, *31*(5), 685 - 712. Retrieved from `http://www.sciencedirect.com/science/article/pii/0749596X9290035V`

Anderson, A. J., Bruni, E., Bordignon, U., Poesio, M., & Baroni, M. (2013). Of words, eyes and brains: Correlating image-based distributional semantic models with neural representations of concepts. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1960–1970). Seattle, Washington, USA: Association for Computational Linguistics. Retrieved from `http://www.aclweb.org/anthology/D13-1202`

Anderson, A. J., Bruni, E., Lopopolo, A., Poesio, M., & Baroni, M. (2015). Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *NeuroImage*, *120*, 309-322.

Baayen, H. R., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database. release 2 (CD-ROM)*. Philadelphia, Pennsylvania: Linguistic Data Consortium, University of Pennsylvania.

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the n400. *Language and Cognitive Processes*, *26*(9), 1338-1367. Retrieved from `https://doi.org/10.1080/01690965.2010.542671`

Bar, M. (2011). *Predictions in the brain: using our past to generate a future*. Oxford University Press.

Baron, S. G., & Osherson, D. N. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *NeuroImage*, *55*, 1847-1852.

Baron, S. G., Thompson-Schill, S. L., Weber, M., & Osherson, D. (2010). An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. *Cognitive Neuroscience*, *1*(1), 44-51. Retrieved from https://doi.org/10.1080/17588920903548751 (PMID: 24168244)

Baroni, M., & Lenci, A. (2011). How we BLESSed distributional semantic evaluation. In *Proceedings of the 2011 conference on empirical methods in natural language processing* (pp. 1–10"). Edinburgh, UK: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/W11-2501

Bemis, D. K., & Pylkkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, *31*(8), 2801–2814. Retrieved from https://www.jneurosci.org/content/31/8/2801

Bemis, D. K., & Pylkkänen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral cortex*, *23 8*, 1859-73.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience*, *4*(4), 417. Retrieved from +http://dx.doi.org/10.1093/scan/nsp053

Berkum, J. J. A. V., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the n400. *Journal of Cognitive Neuroscience*, *11*, 657-671.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*, 527-536.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*(12), 2767-2796. Retrieved from https://doi.org/10.1093/cercor/bhp055

Blanchard, H. E., & Iran-Nejad, A. (1987). Comprehension processes and eye movement patterns in the reading of surprise-ending stories. *Discourse Processes*, *10*(1), 127-138. Retrieved from https://doi.org/10.1080/01638538709544663

Bonhage, C. E., Fiebach, C. J., Bahlmann, J., & Mueller, J. L. (2014). Brain signature of working memory for sentence structure: Enriched encoding and facilitated maintenance. *Journal of Cognitive Neuroscience*, *26*, 1654-1671.

Booth, R. W., & Weger, U. W. (2013). The function of regressions in reading: Backward eye movements allow rereading. *Memory & Cognition*, *41*(1), 82–97. Retrieved from https://doi.org/10.3758/s13421-012-0244-y

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2012). Linguistic sequence processing and the prefrontal cortex. *The Open Medical Imaging Journal*, *6*(1-M2), 47 - 61.

Borsley, R. D. (1998). *Syntactic theory :a unified approach* (2. ed. ed.). London: E. Arnold.

(Previous ed.: 1991)

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam sentence corpus. *Journal of Eye Movement Research*, *2*(1), 1–12. Retrieved from `https://bop.unibe.ch/JEMR/article/view/2255`

Brennan, J. R., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*, *120*(2), 163–173.

Brennan, J. R., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, *157–158*, 81 - 94. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0093934X15300687`

Bright, P., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2005). The anatomy of object processing: The role of anteromedial temporal cortex. *The Quarterly Journal of Experimental Psychology Section B*, *58*(3-4b), 361-377. Retrieved from `https://doi.org/10.1080/02724990544000013`

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*.

Busser, B., Daelemans, W., & van den Bosch, A. (1999). Machine learning of word pronunciation: the case against abstraction. In *Sixth european conference on speech communication and technology, EUROSPEECH 1999, budapest, hungary, september 5-9, 1999.* Retrieved from `http://www.isca-speech.org/archive/eurospeech_1999/e99_2123.html`

Caramazza, A. (1996). The brain's dictionary. *Nature*, *380*(485), 1476 - 4687.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*(4), 572 - 582. Retrieved from `http://www.sciencedirect.com/science/article/pii/0093934X76900481`

Castelhano, M., & Rayner, K. (2008). Eye movements during reading, visual search, and scene perception: An overview. In *Cognitive and cultural influences on eye movements* (p. 175 - 195). Tianjin People's Publishing House.

Che, W., Liu, Y., Wang, Y., Zheng, B., & Liu, T. (2018). Towards better UD parsing: Deep contextualized word embeddings, ensemble, and treebank concatenation. In *Proceedings of the CoNLL 2018 shared task: Multilingual parsing from raw text to universal dependencies* (pp. 55–64). Brussels, Belgium: Association for Computational Linguistics.

Retrieved from http://www.aclweb.org/anthology/K18-2005

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton and Co.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: The MIT Press. Retrieved from http://www.amazon.com/Aspects-Theory-Syntax-Noam-Chomsky/dp/0262530074

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Cohen, L. (2002). Language-specific tuning of visual cortex? functional properties of the visual word form area. *Brain*, *125*(5), 1054–1069.

Daelemans, W., & van Den Bosch, A. (2005). Memory-based learning in natural language processing. In *Memory-based language processing* (p. 3–14).

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499-505.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193 - 210. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010027708001741

D'Esposito, M., & Postle, B. R. (2007). The cognitive neuroscience of working memory. *Annual review of psychology*, *66*, 115-42.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

de Vries, C., Reijnierse, W. G., & Willems, R. M. (2018). Eye movements reveal readers' sensitivity to deliberate metaphors during narrative reading [Journal Article]. *Scientific Study of Literature*, *8*(1), 135-164. Retrieved from https://www.jbe-platform.com/content/journals/10.1075/ssol.18008.vri

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. J., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1–2), 145 - 177. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010027703002300

Faber, M., Mak, M., & Willems, R. (2020). Word skipping as an indicator of individual reading style during literary reading. *Journal of Eye Movement Research*, *13*(3). Retrieved from https://bop.unibe.ch/JEMR/article/view/JEMR.13.3.2

Fares, M., Kutuzov, A., Oepen, S., & Velldal, E. (2017). Word vectors, reuse, and replicability: Towards a community repository of large-text resources. In *Proceedings of the 21st nordic conference on computational linguistics* (pp. 271–276). Gothenburg, Sweden: Association for Computational Linguistics. Retrieved from http://www.aclweb

.org/anthology/W17-0237

Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*(5), 475-494.

Frank, S. L., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, *22*, 829-834.

Frank, S. L., Bod, R., & Christiansen, M. H. (2012). How hierarchical is language use? *Proceedings of the Royal Society B: Biological Sciences*, *279*, 4522–4531.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain Lang*, *140*, 1-11.

Frank, S. L., & Thompson, R. L. (2012). Early effects of word surprisal on pupil size during reading. In *Cognitive science society* (pp. 1554–1559).

Frank, S. L., & Willems, R. M. (2017). Word predictability and semantic similarity show distinct patterns of brain activity during language comprehension. *Language, Cognition and Neuroscience*, *32*(9), 1192-1203. Retrieved from https://doi.org/10.1080/23273798.2017.1323109

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178 - 210. Retrieved from http://www.sciencedirect.com/science/article/pii/0010028582900081

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78 - 84. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364661300018398

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357-1392. Retrieved from https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I., & von Cramon, D. Y. (2005). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral cortex*, *16 12*, 1709-17.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, *74*, 289-300.

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex*, *13*(2), 170. Retrieved from +http://dx.doi.org/10.1093/cercor/13.2.170

Friston, K. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221.

Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, *1*(1), 13–36.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, *6*(3), 218–229.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., & Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Current Biology*, *7*(9), 645 - 651. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0960982206002910`

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*(12), 474 - 480. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1364661308002222`

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M. (1997). *Algemene Nederlandse Spraakkunst.* ONijhoff and Deurne: Wolters Plantyn.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416 - 423. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1364661305002093`

Hagoort, P. (2013a). MUC (Memory, Unification, Control) and beyond. *Front Psychology*, *4*, 416.

Hagoort, P. (2013b). MUC (Memory, Unification, Control) and beyond. *Front Psychology*, *4*, 416.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences,* 4^{th} *edition* (p. 819-836). Cambridge (MA): MIT Press.

Hagoort, P., & Levelt, W. J. M. (2009). The speaking brain. *Science*, *326*(5951), 372–373. Retrieved from `https://science.sciencemag.org/content/326/5951/372`

Hagoort, P., & van Berkum, J. J. (2007). Beyond the sentence given. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *362 1481*, 801-811.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. *Second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies 2001 - NAACL '01.*

Hassan, A., & Mahmood, A. (2017). Deep learning for sentence classification. *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, 1-5.

Herrmann, B., Obleser, J., Kalberlah, C., Haynes, J.-D., & Friederici, A. D. (2012). Dissociable neural imprints of perception and grammar in auditory functional imaging. *Human Brain Mapping*, *33*(3), 584–595. Retrieved from http://dx.doi.org/10.1002/hbm.21235

Hertrich, I., Dietrich, S., & Hermann, A. (2016). The role of the supplementary motor area for speech and language processing. *Neurosci Biobehav Rev.*, *Sep*(68), 602-610.

Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organisation of semantic memory. *Memory*, *3*(3-4), 463-495. Retrieved from https://doi.org/10.1080/09658219508253161 (PMID: 8574874)

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic Dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*(6), 1783-1806. Retrieved from https://doi.org/10.1093/brain/115.6.1783

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665-679. Retrieved from https://doi.org/10.1162/jocn.2006.18.4.665

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *NeuroImage*, *36*, 924-932.

Hyönä, J. (1995). An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology Learning Memory and Cognition*, *21*, 1365-1373.

Indefrey, P., & Cutler, A. (2004). Prelexical and lexical processing in listening. In M. Gazzaniga (Ed.), *The cognitive neurosciences iii* (p. 759-774).

Jackendoff, R. (2003). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press. Retrieved from https://books.google.nl/books?id=228WWikKa-UC

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research*, *1146*, 2–22.

Jackendoff, R., & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In *Measuring grammatical complexity* (p. 65 — 82). Oxford: Oxford University Press.

Kemmerer, D. L. (2014). *Cognitive neuroscience of language*. Psychology Press. Retrieved from https://books.google.nl/books?id=nYi9oAEACAAJ

Kennedy, A. (1992). The spatial code hypothesis. In K. Rayner (Ed.), *Eye movements and visual cognition. springer series in neuropsychology.* New York: Springer.

Keuleers, E., Brysbaert, M., & New, B. (2010). Subtlex-nl: A new measure for dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650.

Klerke, S., Barrett, M., Castilho, S., & Søgaard, A. (2015). Reading metrics for estimating task efficiency with SMT output. In *Proceedings of the sixth workshop on cognitive aspects of computational language learning* (pp. 6–13). Association for Computational Linguistics.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). *Annual review of psychology*, *62*, 621-647.

Lam, N. H. L., Schoffelen, J.-M., Uddén, J., Hultén, A., & Hagoort, P. (2016). Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *NeuroImage*, *142*, 43-54.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Li, J., & Hale, J. (2019). Grammatical predictors for fMRI time-courses. In R. C. Berwick & E. P. Stabler (Eds.), *Minimalist parsing* (p. 159-173). Oxford: Oxford University Press.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, *13*(3), 361 - 392. Retrieved from `http://www.sciencedirect.com/science/article/pii/001002778390015X`

Lopopolo, A., Frank, S. L., van den Bosch, A., & Willems, R. M. (2017). Using stochastic language models (slm) to map lexical, syntactic, and phonological information processing in the brain. *PLOS ONE*, *12*(5), 1-18. Retrieved from `https://doi.org/10.1371/journal.pone.0177794`

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, *126*(2), 313–318.

Mak, M., & Willems, R. M. (2018). Mental simulation during literary reading: Individual differences revealed with eye-tracking. *Language, Cognition and Neuroscience*, *34*(4), 511–535. Retrieved from https://doi.org/10.1080/23273798.2018.1552007

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the ifg and psts in syntactic prediction: Evidence from a parametric study of hierarchical structure in fmri. *Cortex*, *88*, 106 - 123. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010945216303549

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., ... Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467-479. Retrieved from https://doi.org/10.1162/jocn.1993.5.4.467 (PMID: 23964919)

McCann, B., Bradbury, J., Xiong, C., & Socher, R. (2017). Learned in translation: Contextualized word vectors. In *Nips.*

Melamud, O., Goldberger, J., & Dagan, I. (2016). context2vec: Learning generic context embedding with bidirectional lstm. In *Conll.*

Mel'čuk, I. (1988). *Dependency syntax: Theory and practice*. State University of New York Press.

Mendoza, J. E. (2011). Heschl's gyrus. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 1243–1243). New York, NY: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-0-387-79948-3_737

Mikolov, T., Chen, K., Corrado, G. S., & Dean, J. (2013). Efficient estimation of word representations in vector space. *CoRR*, *abs/1301.3781*.

Mitchell, T. M., Shinkareva, S. V., Carlson, A., min Chang, K., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, *320 5880*, 1191-5.

Monsalve, I. F., Frank, S. L., & Vigliocco, G. (2012). Lexical surprisal as a general predictor of reading time. In *Proceedings of the 13th conference of the european chapter of the association for computational linguistics* (pp. 398–408). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2380816.2380866

Moss, H. E., Rodd, J. M., Stamatakis, E. A., Bright, P., & Tyler, L. K. (2004). Anteromedial Temporal Cortex Supports Fine-grained Differentiation among Objects. *Cerebral Cortex*, *15*(5), 616-627. Retrieved from https://doi.org/10.1093/cercor/bhh163

Mummery, C. J., Patterson, K., Veltman, D. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Annals of neurology*, *47 1*, 36-45.

Nastase, S., Iacovella, V., & Hasson, U. (2013). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Human Brain Mapping*, *35*(4), 1111–1128.

Nee, D. E., & D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *eLife*, *5*.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., … Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*(18), E3669–E3678. Retrieved from `https://www.pnas.org/content/early/2017/04/11/1701590114`

Nichols, T., Brett, M., Andersson, J., Wager, T., & Polinee, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*, 653-660.

Nijhof, A. D., & Willems, R. M. (2015). Simulating fiction: Individual differences in literature comprehension revealed with fmri. *PLOS ONE*, *10*, 1-17. Retrieved from `https://doi.org/10.1371/journal.pone.0116492`

Nivre, J., & Kübler, S. (2009). Dependency parsing. *Synthesis Lectures on Human Language Technologies*, *2*.

Noord, G. V. (2006). At last parsing is now operational. In *In taln 2006* (pp. 20–42).

Obleser, J., Meyer, L., & Friederici, A. D. (2011). Dynamic assignment of neural resources in auditory comprehension of complex sentences. *NeuroImage*, *56*(4), 2310–2320.

Oostdijk, N. (2000). The spoken dutch corpus. overview and first evaluation. In *Proceedings of the second international conference on language resources and evaluation (lrec-2000).* Athens, Greece: European Language Resources Association (ELRA). Retrieved from `http://www.lrec-conf.org/proceedings/lrec2000/pdf/110.pdf`

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *463*, 461-488.

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604-609. Retrieved from `https://doi.org/10.1093/scan/nss055`

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, *108*(6), 2522–2527. Retrieved from `https://www.pnas.org/content/108/6/2522`

Parviz, M., Johnson, M., Johnson, B., & Brock, J. (2011). Using language models and latent semantic analysis to characterise the n400m neural response. In *Proceedings of the australasian language technology association workshop 2011* (pp. 38–46). Canberra, Australia. Retrieved from https://www.aclweb.org/anthology/U11-1007

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? the representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*, 976-987.

Peters, M. E., Ammar, W., Bhagavatula, C., & Power, R. (2017). Semi-supervised sequence tagging with bidirectional language models. In *Acl*.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. In *Proc. of naacl*.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Poser, B. A., Koopmans, P. J., Witzel, T., Wald, L. L., & Barth, M. (2010). Three dimensional echo-planar imaging at 7 Tesla. *NeuroImage*, *51*(1), 261 - 266. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811910001473

Price, C., & Devlin, J. (2003). The myth of the visual word form area. *Neuroimage*, *19*(3), 473-481.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Rayner, K., & Pollatsek, A. (1995). *The Psychology of Reading*. Hillsdale: Lawrence Erlbaum Associates.

Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition*, *22*(3), 281–285. Retrieved from https://doi.org/10.3758/BF03200855

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using E–Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110 - 149. Retrieved from http://www.sciencedirect.com/science/article/pii/S0273229713000129

Rogalsky, C., & Hickok, G. (2008). Selective Attention to Semantic and Syntactic Features Modulates Sentence Processing Networks in Anterior Temporal Cortex. *Cerebral Cortex*, *19*(4), 786-796. Retrieved from https://doi.org/10.1093/cercor/bhn126

Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(3), 201–213. Retrieved from `https://doi.org/10.3758/CABN.6.3.201`

Rogers, T. T., Ralph, M. A. L., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological review*, *111 1*, 205-35.

Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, *111*(1), 1 - 7. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0093934X09000418`

Schäfer, R., & Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In N. Calzolari et al. (Eds.), *Lrec* (p. 486-493). European Language Resources Association (ELRA). Retrieved from `http://dblp.uni-trier.de/db/conf/lrec/lrec2012.html#SchaferB12`

Schoffelen, J.-M., Hultén, A., Lam, N., Marquand, A. F., Uddén, J., & Hagoort, P. (2017). Frequency-specific directed interactions in the human brain network for language. *Proceedings of the National Academy of Sciences*.

Schoffelen, J.-M., Oostenveld, R., Lam, N., Udden, J., Hultén, A., & Hagoort, P. (2019). A 204-subject multimodal neuroimaging dataset to study language processing. *Scientific Data*, *6*.

Seuren, P. (2009). *Language from within: Vol. 1. language in cognition*. Oxford, UK: Oxford University Press.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J. J. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research*, *17*(1), 75 - 82. Retrieved from `http://www.sciencedirect.com/science/article/pii/S092664100300082X`

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319.

Snijders, T. M., Vosse, T., Kempen, G., Berkum, J. J. A. V., Petersson, K. M., & Hagoort, P. (2008). Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*(7), 1493–1503.

Spivey, M., Richardson, D., & A. Fitneva, S. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. In J. Henderson & F. Ferreira (Eds.), *Interfacing language, vision, and action* (p. 161-190). San Diego: CA: Academic Press.

Stabler, E. (1997). Derivational minimalism. In C. Retoré (Ed.), *Logical aspects of computational linguistics* (pp. 68–95). Berlin, Heidelberg: Springer Berlin Heidelberg.

Steen, G. (2008). The paradox of metaphor: Why we need a three-dimensional model of metaphor. *Metaphor and Symbol*, *23*(4), 213-241. Retrieved from https://doi.org/10.1080/10926480802426753

Stolcke, A. (2002). Srilm - an extensible language modeling toolkit. In (pp. 901–904).

Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., & Zwarts, F. (1998). Localizing components of a complex task: sentence processing and working memory. *Neuroreport*, *9 13*, 2995-2999.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409.

Tesnière, L., Osborne, T. J., & Kahane, S. (2015). *Elements of structural syntax*. John Benjamins Publishing Company. Retrieved from https://books.google.nl/books?id=FNjooAEACAAJ

Thompson-Schill, S., Ramscar, M., & Chrysikou, E. (2009). Cognition Without Control: When a Little Frontal Lobe Goes a Long Way. *Current Directions in Psychological Science*, *18*(5), 259–263. Retrieved from http://dx.doi.org/10.1111/j.1467-8721.2009.01648.x

Thothathiri, M., Kimberg, D. Y., & Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *Journal of Cognitive Neuroscience*, *24*(1), 212-222. (PMID: 21861679)

Tremblay, P., Deschamps, I., Baroni, M., & Hasson, U. (2016). Neural sensitivity to syllable frequency and mutual information in speech perception and production. *NeuroImage*, *136*, 106 - 121. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811916301380

Tremblay, P., & Small, S. L. (2011). On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *NeuroImage*, *57*(4), 1561 - 1571. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811911005751

Tyler, L. K., Randall, B., & Stamatakis, E. A. (2008). Cortical differentiation for nouns and verbs depends on grammatical markers. *Journal of Cognitive Neuroscience*, *20*(8), 1381-1389. Retrieved from https://doi.org/10.1162/jocn.2008.20095 (PMID: 18303983)

Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., & Moss, H. E. (2004). Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience*, *16*, 351-362.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage*, *15*(1), 273 - 289. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811901909784

Ullman, M. T. (2012). The biocognition of the mental lexicon. In *The oxford handbook of psycholinguistics.* Oxford University Press, Oxford.

van Essen, R. (2014). *Hier wonen ook mensen*. Atlas Contact, Uitgeverij. Retrieved from https://books.google.nl/books?id=N55HBAAAQBAJ

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering*, *44*(9), 867–880.

Vigneau, M., Beaucousin, V., Hervé, P.-Y., Duffau, H., Crivello, F., Houdé, O., … Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414 - 1432. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811905024511

Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. *Cognitive Processes in Eye Guidance*, 1-32.

Westerlund, M., Kastner, I., Kaabi, M. A., & Pylkkänen, L. (2015). The latl as locus of composition: MEG evidence from english and arabic. *Brain and Language*, *141*, 124 - 134. Retrieved from http://www.sciencedirect.com/science/article/pii/S0093934X14001990

Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia*, *57*, 59-70.

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. In *Front. psychology.*

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & Bosch, A. V. D. (2016). Prediction during natural language comprehension. *Cerebral Cortex*, *26*(6), 2506–2516.

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and language*, *101*(3), 278-289.

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, *91*, 412 - 419.

Wulfeck, B., & Bates, E. (1991). Differential sensitivity to errors of agreement and word order in Broca's aphasia. *Journal of Cognitive Neuroscience*, *3*(3), 258-272. Retrieved from https://doi.org/10.1162/jocn.1991.3.3.258 (PMID: 23964841)

Zaccarella, E., & Friederici, A. D. (2015). Merge in the human brain: A sub-region based functional investigation in the left pars opercularis. *Frontiers in Psychol-*

ogy, *6*, 1818. Retrieved from https://www.frontiersin.org/article/10.3389/fpsyg.2015.01818

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2015). Building by Syntax: The Neural Basis of Minimal Linguistic Structures. *Cerebral Cortex*, *27*(1), 411-421. Retrieved from https://doi.org/10.1093/cercor/bhv234

Curriculum Vitæ

Areas of specialization

Cognitive neuroscience of language, Computational linguistics, Neural decoding, MVPA, Language processing in humans and machines, Semantic-based brain computer interface for communication, Multimodal distributional semantics, Image-based distributional semantics, Sound-based distributional semantics, Cognitive and Neural modelling

Education

29-09-2019 - 28-12-2019: Visiting researcher at the Machine Learning Department of the Carnegie Mellon University, Pittsburgh, PA, USA.

2016-2020: PhD Candidate in the Language in Interaction Consortium at the Center for Language Studies of the Radboud University, Nijmegen, NL.
Title of the project: "Encoding and decoding the neural signatures of natural language comprehension", supervisors: prof. dr. A. van den Bosch, co-supervisors: dr. K. M. Petersson, dr. R. Willems and dr. S. Frank.

2011-2014: MSc Cum Laude *Cognitive Science (Language and Multimodal Interaction track)*, University of Trento, IT.
Thesis: "Words as abstract or perceptual symbols, experimental insights from artificial intelligence and neuroscience", supervisor: prof. dr. Marco Baroni, co-supervisor: dr. Angelika Lignau.

2009-2011: MA *Linguistics*, University of Amsterdam, NL.
Thesis: "The influence of population structure and acquisition dynamics on language complexity, simulating the variation in complexity and population structure using the Language Game

model", supervisor: dr. Bart de Boer, co-supervisor: prof. dr. Paul Boersma;

2004-2008: BA Cum Laude *General Linguistics and Modern Languages*, University of Bologna, IT;

Grants, honors & awards

12-02-2015: Premio di Merito (Graduate Merit Award), University of Trento, IT;

2006-2007: Socrates/Erasmus Scholarship, University of Bologna - Trinity College Dublin, IE.

Publications

Papers

2020: Lopopolo, Alessandro, Antal van den Bosch, Karl-Magnus Petersson and Roel M. Willems. Distinguishing syntactic operations in the brain: Dependency and phrase-structure parsing. [submitted];

2020: Lopopolo, Alessandro, Antal van den Bosch, Jan-Mathijs Schoffelen and Roel M. Willems. Words in context: tracking context-processing during language comprehension using computational language models and MEG. [submitted];

2019: Lopopolo, Alessandro, Stefan L. Frank, Antal van den Bosch and Roel M. Willems. Dependency Parsing with your Eyes: Dependency Structure Predicts Eye Regressions During Reading. Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics 2019. NAACL 2019;

2018: Lopopolo, Alessandro Stefan L. Frank, Antal van den Bosch, Anabel Nijhof and Roel M. Willems. The Narrative Brain Dataset (NBD), an fMRI Dataset for the Study of Natural Language Processing in the Brain. Proceedings of the Eleventh International Conference on Language Resources and Evaluation 2018. LREC 2018;

2017: Lopopolo, Alessandro Stefan L. Frank, Antal van den Bosch and Roel M. Willems. Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain. PLOS ONE 12(5);

2015: Anderson, Andrew James, Elia Bruni, Alessandro Lopopolo, Massimo Poesio and Marco Baroni. Reading visually embodied meaning from the brain: visually grounded computational models decode visual-object mental imagery induced by written text. NeuroImage, Volume 120, 15 October 2015, Pages 309-322;

2015: Lopopolo, Alessandro and Emiel van Miltenburg. 2015. Sound-based distributional models. In Proceedings of the 11th International Conference on Computational Semantics, pages 70-75, London, UK, April 2015. Association for Computational Linguistics;

2015: Vossen, Piek, Tommaso Caselli, Filip Ilievski, Ruben Izquierdo, Alessandro Lopopolo, Emiel van Miltenburg, Roser Morante, Minh Ngoc Le, and Marten Postma. Words in context: a reference perspective on the lexicon. In Proceedings of the Workshop on Multiple Approaches for Multilingual frame Semantics (MAPLEX-2015), Yamagata, Japan.

2011: Lopopolo, Alessandro and Tamas Biro. Language Change and SAOT, the case of sentential negation. CLIN Journal vol. 1.

Samenvatting

In dit proefschrift heb ik onderzocht of taalverwerking in de hersenen kan worden onderverdeeld in verschillende subcomponenten. Hiervoor heb ik computationele taalmodellen en neurale gegevens die zijn verzameld tijdens een naturalistisch stimulusparadigma gebruikt.

Tijdens taalverwerking ontvangen de hersenen een in wezen sequentieel signaal. De input ontvouwt zich immers lineair in tijd of ruimte. Niettemin zijn slechts van spraakgeluiden of grafemen niet voldoende voor taalverwerking. Het primaire doel van dit proefschrift is de diversiteit aan informatietypen, structuren, en bewerkingen die deel uitmaken van natuurlijke taalverwerking, te onderzoeken. De eerste intuïtie is dat de hersenen geen eenvoudigweg opeenvolgende verwerking gebruiken waarin het ene type representatie na het andere wordt berekend. Verwerking bestaat daarentegen uit een veelvoud van naast elkaar bestaande en mogelijk synchrone en bovenop elkaar werkende processen. In dit proefschrift zal ik me concentreren op de sequentiële eigenschappen van de stimulus, het type syntactische structuren dat door de hersenen wordt berekend, en de basishandelingen die de verwerking ondersteunen.

De **sequentiële eigenschappen van de stimulus** bestaan uit de informatie die de hersenen uit de stimulus zelf halen. In dit proefschrift worden deze eigenschappen geanalyseerd op het niveau van fonemen, woorden en grammaticale klassen. De volgende zin: *de jongen leest een boek in bed*, kan worden opgevat als een reeks woorden (*de*, *jongen*, enz.), Maar ook als een reeks grammaticale categorieën (*artikel*, *naamwoord*, *werkwoord*, *artikel*, *naamwoord*, *voorzetsel*, *naamwoord*), of als een reeks fonemen. Het is de vraag of het menselijk brein de zin verwerkt als een reeks woorden, als een reeks grammaticale categorieën, en als een reeks fonemen op dezelfde manier en in dezelfde hersenstructuren. Deze verschillende niveaus van sequentiële informatie worden gemodelleerd met behulp van stochastische taalmodellen, die de sequenties beschrijven in termen van voorwaardelijke waarschijnlijkheden. Dat zijn modellen die de kans berekenen dat het woord *boek* volgt na *...leest een*, of dat een *naamwoord* volgt op een *werkwoord*.

Naast deze sequentiële eigenschappen zullen de hersenen waarschijnlijk ook de **syntactische structuren** van de taalinvoer berekenen. Met andere woorden, de hersenen bouwen een netwerk op van relaties tussen de lexicale eenheden die de inkomende stimulus vormen. Het brein herkent bijvoorbeeld een subject-predikaatrelatie tussen de woorden 'jongen' en 'leest'. In dit proefschrift beschouw ik twee soorten syntaxis: één gebaseerd op **dependency grammars** en één gebaseerd op **phrase-structure grammars**. De vraag is of het brein syntactische structuren opbouwt volgens principes die meer in lijn zijn met de ene of de andere grammatica.

Ten slotte bedoel ik met **basisbewerkingen** het ophalen van lexicale informatie uit het mentale lexicon en hun integratie in grotere representaties, zoals een zin. De intuïtie is dat tijdens het begrijpen van een zin ons brein de betekenis van elk woord toewijst door te putten uit zijn eigen mentale lexicon en dat deze vervolgens worden gecombineerd tot geleidelijk grotere structuren. Om dit te bestuderen, gebruik ik computationele representaties van de betekenis van woorden verkregen uit kunstmatige neurale netwerken (deep learning word embeddings).

De resultaten tonen aan dat verschillende soorten sequentiële informatie een rol spelen in verschillende hersengebieden die tot het taalnetwerk van de hersenen behoren. Bovendien lijken verschillende soorten syntactische grammatica's te worden verwerkt door verschillende hersengebieden die gevoelig zijn voor syntaxis. Ten slotte kan de integratie van lexicale informatie worden gemodelleerd met behulp van state-of-the-art computational word embeddings.

Summary

In this thesis, I have explored the hypothesis that language processing in the brain can be divided into different sub-components using different computational linguistic models and neural data collected during a naturalistic stimulus paradigm.

During language processing, the brain receives a signal that is essentially sequential. The input unfolds linearly in time or space. Nevertheless, arrays of speech sounds or graphemes are not all there is to support language processing. The primary aim of this thesis is to investigate the diversity of information types, structures and operations that are part of natural language processing. The first intuition is that the brain does not support a simple sequential pipeline in which one type of information is processed at the same time and one type of representation is calculated after another. Processing consists of a multitude of coexisting and possibly synchronous and superimposed processes. I will focus on the sequential properties of the stimulus, the type of syntactic structures computed by the brain, and the basic operations underpinning processing.

The **sequential properties of the stimulus** consist of the information that the brain extracts directly from the stimulus. In this thesis, these properties are analyzed at the level of phonemes, words and grammatical categories. The following sentence: *the boy is reading a book in bed*, can be understood as a series of words (*the*, *boy*, etc.), but also as a series of grammatical categories (*article*, *noun*, *verb*, *article*, *noun*, *preposition*, *noun*), or as a sequence of phonemes. The question is whether the human brain processes the sentence as a series of words, as a series of grammatical categories, and as a series of phonemes in the same way and in the same brain structures. These different levels of sequential information are modeled using stochastic language models, which describe the sequences in terms of conditional probabilities. Such models calculate the probability of hearing the word *book* after having heard ... *reads a*, or encountering a *noun* after a *verb*.

In addition to these sequential features, the brain is also likely to analyze the **syntactic structures** of the linguistic input. In other words, the brain builds a network of relationships between

the lexical units that make up the incoming stimulus. For example, the brain recognizes the existence of a subject-predicate relationship between the words *boy* and *reads*. In this thesis, I consider two types of syntactic formalisms: one based on **dependency grammars** and one based on **phrase-structure grammars**. The question is whether the brain builds syntactic structures according to principles that are more in line with one or the other grammar.

Finally, by **basic operations** I mean the recovery of lexical information from the mental lexicon and its integration into larger representations. The intuition is that while understanding a sentence, our brain assigns the meaning of each word by drawing on its own mental lexicon and then combines these into gradually larger structures. To study this, I use computational representations of the meaning of words obtained from artificial neural networks (deep learning word embeddings).

The results show that different types of sequential information can be detected in different areas of the brain that belong to the brain's language network. In addition, different types of syntactic grammars seem to be processed by different areas of the brain that are sensitive to syntax. Finally, the integration of lexical information can be modeled using state-of-the-art computational word embeddings.

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*

25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*

51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*

98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*
122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*
124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Hoey*
126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*
127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*
128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*
130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*
131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*
132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*
133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*
134. Indirect request comprehension in different contexts. *Johanne Tromp*
135. Envisioning Language - An Exploration of Perceptual Processes in Language Comprehension. *Markus Ostarek*
136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*
137. Let the agents do the talking: on the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*
138. Age and hearing loss effects on speech processing. *Xaver Koch*
139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*
140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*
141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*
142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*
143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas-Berscia*

144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*
145. Understanding temporal overlap between production and comprehension. *Amie Fairs*
146. The role of exemplars in speech comprehension. *Annika Nijveld*
147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*
148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*
149. Reason-giving in everyday activities. *Julija Baranova*
150. Speech planning in dialogue - Psycholinguistic studies of the timing of turn taking. *Mathias Barthel*
151. The role of neural feedback in language unification: How awareness affects combinatorial processing. *Valeria Mongelli*
152. Exploring social biases in language processing. *Sara Iacozza*
153. Vocal learning in the pale spear-nosed bat, Phyllostomus discolor. *Ella Lattenkamp*
154. The effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands. *Elif Zeynep Azar*
155. Language and society: How social pressures shape grammatical structure. *Limor Raviv*
156. The moment in between: Planning speech while listening. *Svetlana-Lito Gerakaki*
157. How speaking fast is like running: Modelling control of speaking rate. *Joe Rodd*
158. The power of context: How linguistic contextual information shapes brain dynamics during sentence processing. *René Terporten*
159. Neurobiological models of sentence processing. *Marvin Uhlmann*
160. Individual differences in syntactic knowledge and processing: The role of literacy experience. *Saoradh Favier*
161. Memory for speaking and listening. *Eirini Zormpa*
162. Masculine generic pronouns: Investigating the processing of an unintended gender cue. *Theresa Redl*
163. Properties, structures and operations: Studies on language processing in the brain using computational linguistics and naturalistic stimuli. *Alessandro Lopopolo*