

RESEARCH ARTICLE

Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain

Alessandro Lopopolo^{1*}, Stefan L. Frank¹, Antal van den Bosch^{1,3}, Roel M. Willems^{1,2,4}

1 Centre for Language Studies, Radboud University Nijmegen, Nijmegen, the Netherlands, **2** Donders Institute, Radboud University Nijmegen, Nijmegen, the Netherlands, **3** Meertens Institute, Royal Netherlands Academy of Science and Arts, Amsterdam, the Netherlands, **4** Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

* a.lopopolo@donders.ru.nl



OPEN ACCESS

Citation: Lopopolo A, Frank SL, van den Bosch A, Willems RM (2017) Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain. PLoS ONE 12(5): e0177794. <https://doi.org/10.1371/journal.pone.0177794>

Editor: Philip Allen, University of Akron, UNITED STATES

Received: January 16, 2017

Accepted: May 3, 2017

Published: May 18, 2017

Copyright: © 2017 Lopopolo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: ***PA at accept: please ask AU if "Other" file called "Related study, data source" is meant to be pubbed. We have made available the whole 24-subject fMRI dataset and related scripts, data and meta-data on <https://osf.io/75tx9/> (DOI [10.17605/OSF.IO/75TX9](https://doi.org/10.17605/OSF.IO/75TX9)).

Funding: The work presented here was funded by Netherlands Organisation for Scientific Research (NWO) Gravitation Grant 024.001.006 to the Language in Interaction Consortium and by NWO Vidi grant (NWO-Vidi 276-89-007).

Abstract

Language comprehension involves the simultaneous processing of information at the phonological, syntactic, and lexical level. We track these three distinct streams of information in the brain by using stochastic measures derived from computational language models to detect neural correlates of phoneme, part-of-speech, and word processing in an fMRI experiment. Probabilistic language models have proven to be useful tools for studying how language is processed as a sequence of symbols unfolding in time. Conditional probabilities between sequences of words are at the basis of probabilistic measures such as surprisal and perplexity which have been successfully used as predictors of several behavioural and neural correlates of sentence processing. Here we computed perplexity from sequences of words and their parts of speech, and their phonemic transcriptions. Brain activity time-locked to each word is regressed on the three model-derived measures. We observe that the brain keeps track of the statistical structure of lexical, syntactic and phonological information in distinct areas.

1 Introduction

“Ze staat stil en kijkt een poosje naar een punt in de verte” (She stands still and looks for a moment at a point in the distance) is a fluent grammatical sentence in Dutch. It consists of a string of symbols that when spoken, unfolds in time. When tokenized, it is composed of 13 word forms, belonging to seven distinct grammatical categories, and 40 phonemes (see Section 2.4). Although a natural language sentence is a linear sequence of surface forms, it allows to be decomposed into different levels of information at the phrasal, word, and sub-word level. These levels of information are what we might call, following traditional linguistic schools: syntax, lexico-semantics, and phonology [1, 2]. 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 [3]. These different levels are probably processed at least partially

Competing interests: The authors have declared that no competing interests exist.

separately by the language faculty [4, 5]. 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. That is, there are dedicated studies investigating syntactic aspects of comprehension, lexico-semantic aspects of comprehension, phonological aspects of comprehension, and so on (the recent textbook by Kemmerer offers a clear overview of the state of the art in each subdiscipline in the cognitive neuroscience of language [6]). Here, we aim to investigate the neural basis of the three streams of information processing during language comprehension simultaneously, within one experiment. In other words, we want to know whether we can fractionate the neural signature of language processing of natural speech in different, parallel streams within one and the same data set.

In order to do so, we fitted characterizations of lexical, phonological and syntactic information to fMRI data collected while participants listened to narratives. Brain activity time-locked to each word is regressed on the three model-derived measures. Stochastic language models computed at the word, part-of-speech and phoneme level are used to operationalize the concept of parallel lexical, syntactic and phonetic streams. We rely on stochastic models because, as detailed in Section 1.1, they have been proven powerful tools for the study of language comprehension. Such an approach goes beyond the typical studies in which one level of processing is studied per experiment.

We begin with a short overview of work in modelling language comprehension using stochastic language models in Section 1.1. In Section 1.2 we introduce the *Streams of processing* framework. In section 2 we present the methodological details of the study, focusing both on the stochastic language models and the computational linguistic techniques employed to annotate and extract lexical, syntactic and phonological information from a common stimulus. Section 2 also provides details on the fMRI data collection paradigm. Section 3 describes the regression analysis employed to delineate the neural correlates of each level of information. Finally, in Sections 4 and 5 we report the results and draw conclusions according to the hypotheses expressed in the previous sections.

1.1 Language processing as a sequential stochastic process

A number of studies [7–11] 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, \dots, x_{t-1}), the human brain assigns a conditional probability $P(x_t|x_1, \dots, x_{t-1})$ 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 behavioural 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 behavioural correlates. It has been found to predict the duration of spoken words, with shorter words being used in less surprising situations [12, 13]. Following Hale [14] and Levy [15], 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 [16–19]. Reading time has also been shown to be correlated to surprisal of the syntactic category (part-of-speech; PoS) of the word being read [20, 21]. Moreover, Monsalve [17] showed that PoS and word surprisal have independent effects on reading times. This behavioural 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 [22, 23]. The fact that surprisal correlates with the amplitude of a classical ERP component related to language comprehension [24] is another source of evidence that stochastic language modelling is a neuro-cognitively valid approximation of sentence comprehension. In a recent paper, Willems and colleagues [25] 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.

1.2 Probabilistic streams

Our starting point is twofold. On one hand, we hypothesize 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 [6]. 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 1.1, we model the processing in these separate streams 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. In a recent study, Nastase and colleagues [26] 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 in 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.

We approach the problem of disentangling phonology, lexical and syntax by using language stimuli which are not explicitly designed to study one of these levels in isolation [22, 25]. Using one stochastic measure computed on three distinct levels of annotations of the same linguistic stimulus, we want to investigate first of all level-specific processing, what we may refer to as streams of information. On the other hand, we are interested in investigating the issue of whether there exists a central supramodal stochastic processor of the brain (what is called modality-independent in [26]) by finding areas that are sensitive to stochastic measures independently from the level of information they have been computed on.

2 Materials and methods

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.

2.1 Participants and stimuli

We re-analysed data from an fMRI study on language comprehension of auditory presented narrative texts [25]. Here we briefly present the data collection procedure, preprocessing, and stimuli employed. Full details can be found in the original paper.

Twenty-four healthy, native speakers of Dutch (8 males; mean age 22.9, 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 by self-report, and all participants were naive with respect to the purpose of the experiment. Written informed consent was obtained in accordance with the Declaration of Helsinki, and the study was approved by the local ethics committee. Participants were paid either in money or in course credit at the end of the study.

Stimuli consisted of three excerpts from three distinct literary novels extracted from the Spoken Dutch Corpus, “Corpus Gesproken Nederlands” (CGN) [27]. The excerpts were spoken at a normal rate, in a quiet room, by female speakers (one speaker per story). Stimulus durations were 3:49 min (622 words), 7:50 min (1,291 words), and 7:48 min (1,131 words). Reversed speech versions of the stories were created with Audacity 2.03 (<http://www.audacityteam.org/>).

2.2 Procedure

The experimental paradigm consisted in passively listening to the three narratives and their reversed versions (for a total of six sessions) inside the MRI scanner. Each story and its reversed speech counterpart were presented following each other. 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 (<https://www.neurobs.com/>). Auditory stimuli were presented through MR-compatible earphones. In order to make sure participants could correctly perceive the stimuli, the actual experimental sessions were preceded by an in-scanner volume test where a fragment from another story with comparable voice and sound quality was presented and the volume was adjusted to the optimal level based on feedback from the participant.

After the scanning session, participants were tested for their memory and comprehension of the stories. The participants were not informed in advance about the test in order to avoid attentional biases during the passive listening to the stories.

2.3 fMRI data acquisition and preprocessing

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 [28], with high temporal resolution (time to

Table 1. The lexical stream is obtained from the simple sequence of word forms in the stimulus presented to the subjects.

| | | | | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|--------|-------|-------|----------|----------|----------|----------|
| 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} |

<https://doi.org/10.1371/journal.pone.0177794.t001>

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.

Preprocessing was performed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>) and Matlab 2010b (<http://www.mathworks.nl/>). 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 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.

2.4 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 we 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, \dots, w_n of symbols w_i belonging to the vocabulary V containing all the word form, as illustrated in Table 1.

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

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

Table 2. The phonological stream is obtained from the phonetic transcription of the words of the stimulus.

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

<https://doi.org/10.1371/journal.pone.0177794.t002>

Table 3. Summary of the types of grammatical categories (POS) and the number of sub-categories used to approximate sequential syntactic information processing.

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

<https://doi.org/10.1371/journal.pone.0177794.t003>

from which the stimuli for our experiments were taken) and comprises 320 tags (see Table 3). Besides 13 base tags, this method explicitly assign 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) [29]. Table 4 contains an example of PoS annotation of the example sentence presented in Table 1 above.

2.5 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, $P(w_t|w_1, \dots, w_{t-1})$ 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} :

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

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} .

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

The dataset from which probabilities $P(w_t|w_1, \dots, w_{t-1})$ are estimated is a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the

Table 4. The same sentence from Tables 1 and 2 annotated with fine-grained grammatical information using the POS tags described above.

| | | | | | | | | | | | | |
|-----|-------|------|----|-------|-----|--------|------|-----|------|----|-----|-------|
| 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 |

<https://doi.org/10.1371/journal.pone.0177794.t004>

Dutch Corpus of Web (NLCOW2012) [30]. For lexical perplexity, each word of the experimental texts is assigned a value computed by SRILM [31].

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 [32].

The phonological information was extracted from the phonemic transcription of each word in the stimulus set. We used a memory-based grapheme phoneme converter [33] trained on CELEX 2 [34]. 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 (<https://ilk.uvt.nl/wopr/>) trained on CELEX 2 [34]. 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 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 [35]. Before the actual analysis, one regressor modelling 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. 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. Three covariates were computed containing each word's word, PoS and phonemic perplexity measures, constituting our regressors of interest modelling the three information streams introduced above. Besides these, log₂-transformed lexical frequency per word was computed using the Subtlex NL corpus [36], log₂-transformed PoS frequency per word was computed using the CGN corpus [27], and log₂-transformed phoneme frequency average per word was computed using CELEX 2 [34]. 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 modelled 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 [37, 38]. The simulations took the amount of autocorrelation in the data into account, as suggested in the literature [37, 38]. The scripts used were taken from (<https://www2.bc.edu/~slotnics/scripts.html>). Table 5 reports the size thresholds for each regressor contrast separately. All reported clusters of size display results significant at the $P < 0.05$ level, corrected for multiple comparisons.

Table 5. Cluster size thresholds for the perplexity-based regressors.

| Region | Cluster size |
|--------------------------|--------------|
| Word-based perplexity | 92 |
| POS-based perplexity | 92 |
| Phoneme-based perplexity | 97 |

<https://doi.org/10.1371/journal.pone.0177794.t005>

3.1 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 we conducted a preliminary analysis consisting in computing their pairwise correlations. Table 6 reports these correlations (Pearson’s *r*). Both 3-gram perplexity (ppl) and 1-gram frequency (freq) computed at lexical, PoS, and phonological level are included.

The correlations between perplexity measures reported in Table 6 are fairly low, even between lexical and PoS perplexity (0.046). These results indicates that there is no strong relation between the regressors we have employed in our 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.

4 Results

In this section we present the results of the analyses conducted using the three perplexity measures as regressors of interest. In supplementary material S2 File we report the results of a similar analysis conducted using frequencies as regressors of interest.

4.1 Lexical stream

Table 7 lists the areas that show significant activity with regard to the word-based perplexity regressor. This network is displayed in Fig 1 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 left anterior superior temporal gyrus.

4.2 Syntactic stream

Fig 2 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

Table 6. Correlation between the stochastic measures used in the analyses.

| | 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 |

<https://doi.org/10.1371/journal.pone.0177794.t006>

Table 7. Significant effects of lexical stream perplexity.

| 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 |

<https://doi.org/10.1371/journal.pone.0177794.t007>

(r-MTS). The bilateral precentral sulcus is also activated. Large portions of the superior frontal gyrus are also sensitive to these regressor. The list of areas with the coordinates of their activation peaks can be found in [Table 8](#).

4.3 Phonological stream

[Fig 3](#) and [Table 9](#) 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

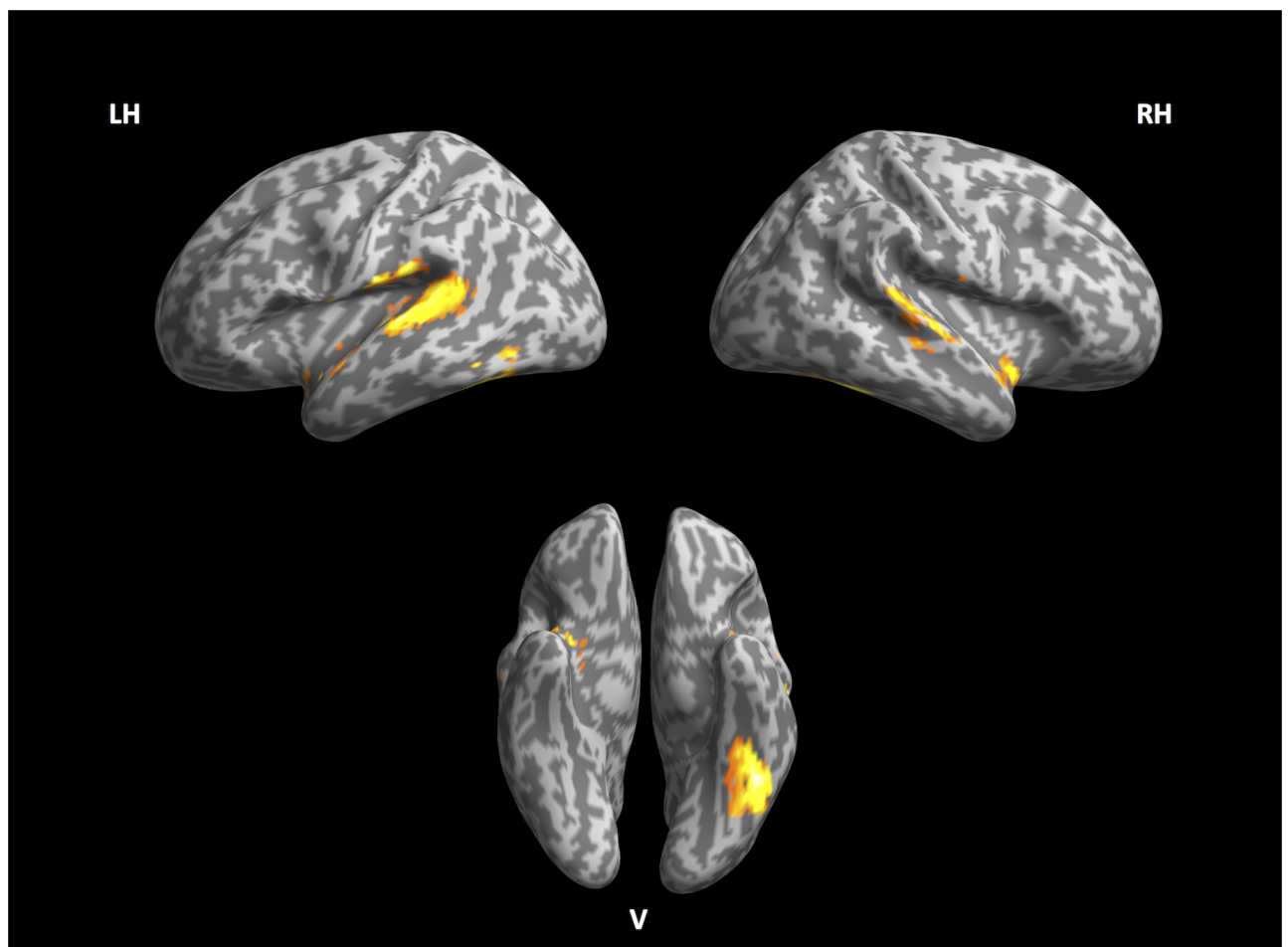


Fig 1. Lexical stream. Left, right and ventral view of inflated cortex plot of the lexical stream.

<https://doi.org/10.1371/journal.pone.0177794.g001>

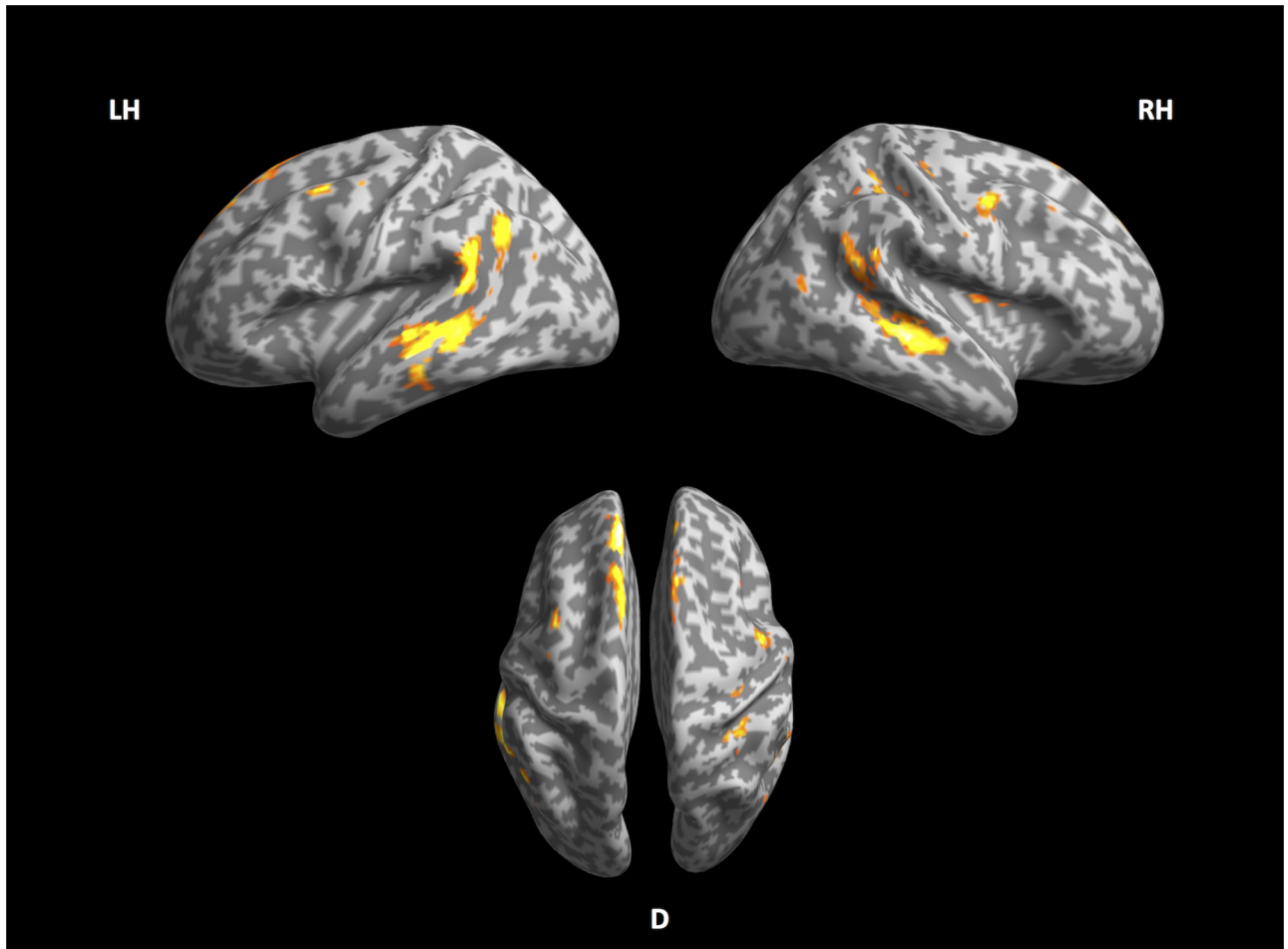


Fig 2. Syntactic stream. Left, right and ventral view of inflated cortex plot of the syntactic stream.

<https://doi.org/10.1371/journal.pone.0177794.g002>

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 (see supplementary material [S1 File](#)). In [S1 File](#) we speculate that this may be caused by the way surprisal and perplexity are computed, and we

Table 8. Significant effects of syntactic stream perplexity.

| 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 |

<https://doi.org/10.1371/journal.pone.0177794.t008>

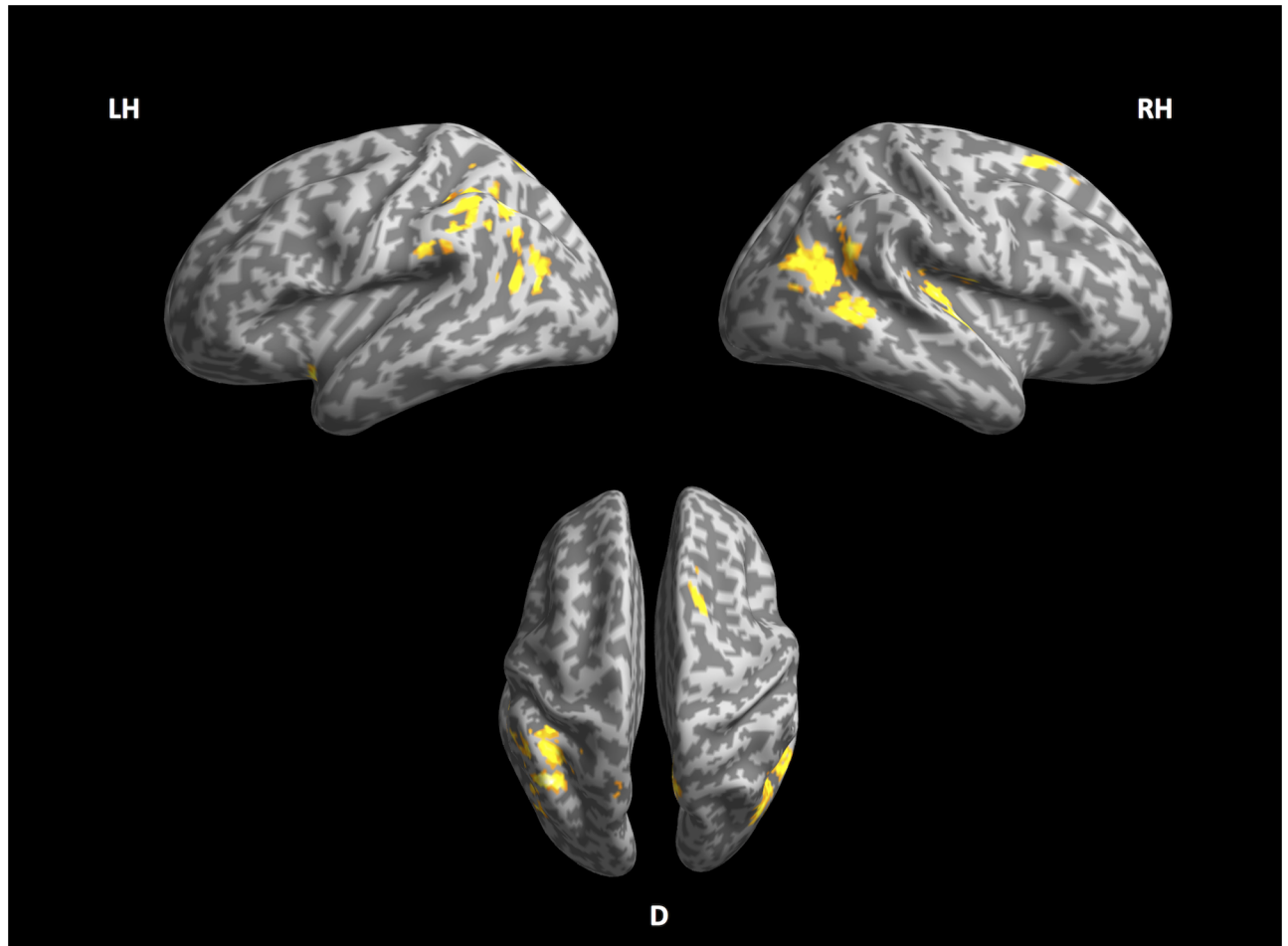


Fig 3. Phonological stream. Left, right and ventral view of inflated cortex plot of the phonological stream.

<https://doi.org/10.1371/journal.pone.0177794.g003>

want to point out this unexpected discrepancy here to highlight that the results of the phonological stream should be interpreted with some more caution than the other two streams.

4.4 Overlap

The results obtained from lexical, syntactic and phonological perplexity allowed us to analyse possible overlap at the cortical level for all three regressor streams. [Table 10](#) contains the name

Table 9. Significant effects of phonological stream perplexity.

| 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 |

<https://doi.org/10.1371/journal.pone.0177794.t009>

Table 10. Areas of overlap between the streams regressors.

| Streams | regions | ~ MNI coordinates |
|--|---------------------------------|-------------------------|
| Syntax \cap Lexical | right STS left posterior MTG | 54 -25 0 -58 -51 11 |
| Syntax \cap Phonology | right AG left AG | 59 -56 29 -53 -64 31 |
| Lexical \cap Phonology | right middle STG | 44 -29 13 |
| Syntax \cap Lexical \cap Phonology | \emptyset | \emptyset |

<https://doi.org/10.1371/journal.pone.0177794.t010>

of regions whose activity is significantly explained by more than one regressor. We 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 [39]. We 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 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 Fig 4.

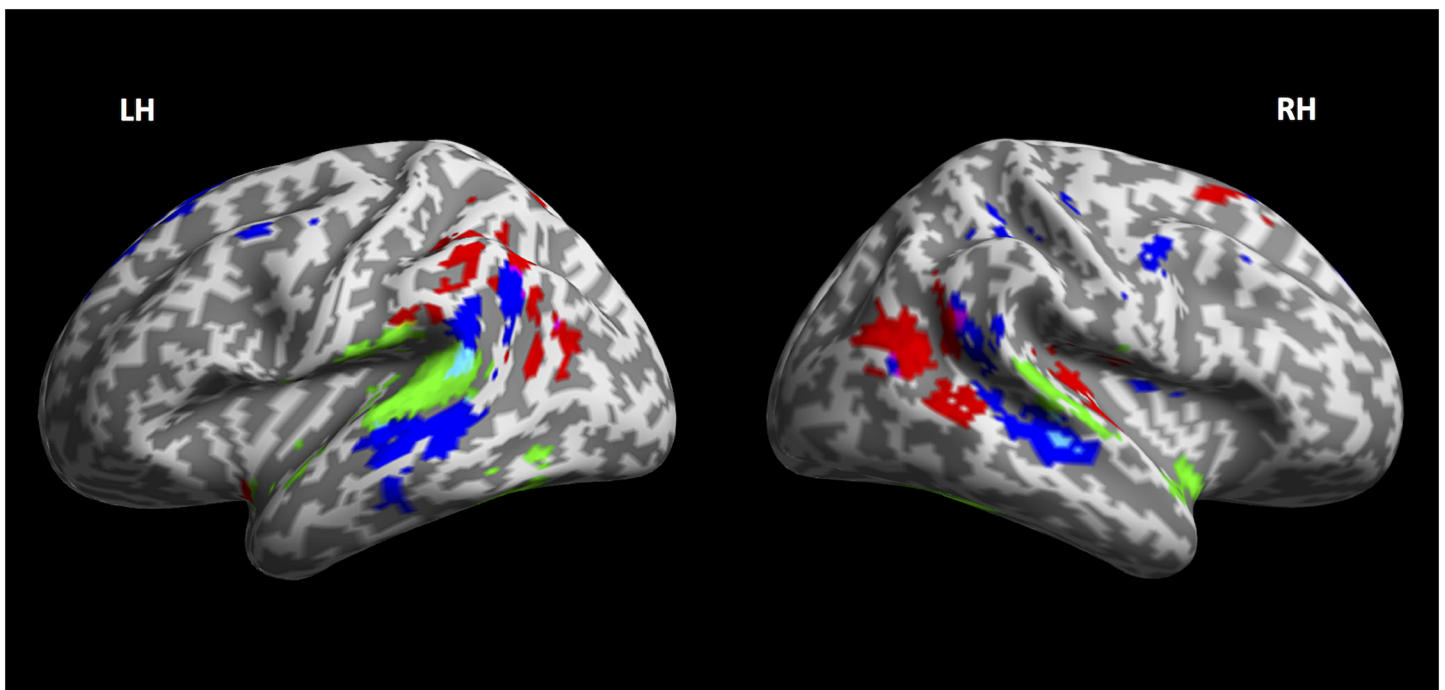


Fig 4. Streams comparison and overlap. Inflated cortex view of the maps of the lexical (green), syntactic (blue) and phonological (red) streams. In this view the overlap between lexical and syntactic streams is particularly evident in the right Middle Temporal Lobe and in the left posterior Superior Temporal Gyrus (cyan). Overlap between syntactic and phonological streams is also evident in the bilateral Angular Gyrus (violet).

<https://doi.org/10.1371/journal.pone.0177794.g004>

5 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.

5.1 Division of labour 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 selecting lexical information and more caudal ones phonological information, with selectivity to PoS-related information in the middle.

5.2 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 [40]. 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 [41], but also in speech processing [42, 43]. 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 and colleagues [44, 45] 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 and colleagues [46] 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 [44].

5.3 Lexical and syntactic streams

5.3.1 Middle temporal gyrus. Studies by Dronkers and colleagues [47] have suggested that the posterior MTG plays a role in retrieving lexical and syntactic properties of incoming words from long-term memory. Hagoort [3] 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 [48]. This intuition is corroborated by our results, which show that activity in this area, although mainly explained by the syntactic regressors employed in our analyses, displays an overlap between the syntactic regressor and the lexical stream. The work of [49] 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 we also observe in our analysis.

Although PoS perplexity intends to model syntactic processing and appears to be a reasonable correlate of the syntactic stream, our analysis shows only a marginal correlation between this regressor and the activity observed in anterior temporal cortex. Studies from Pallier [50], Obleser [51] and Brennan [52] 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 paper. 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 we 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.

5.3.2 Inferior temporal cortex. Activity in the inferior and lateral portions of the left temporal cortex are better explained by the lexical regressor and are likely to be a central hub of the lexical stream. [25] observed the same result using the same dataset, nonetheless interpreting it as activity in the visual word form area (VWFA) [53]. 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 [54] points out that the cortical region corresponding to VWFA is active in normal subjects also during tasks that do not engage visual word form processing. On the other hand, if activity in left Inferior Temporal cortex and specifically in VWFA truly reflects word form prediction, we would have expected also phonological perplexity to show selectivity in this region. Phonological perplexity, computed on the phonemic structure of each 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 [55, 56]. The meta-analysis showed that 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 [57, 58].

5.4 Left inferior frontal gyrus

None of the perplexity-based regressor 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. [59] 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 [60, 61]. Nonetheless, not all results seem to support this view. Brennan and colleagues [52, 62], 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 paper might be due to methodological reasons. Both Brennan and we used naturalistic stimuli and correlation between brain imaging data and stimuli properties (stochastic in our 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 [49], or grammatical vs. ungrammatical [63, 64].

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

6 Conclusions

In this paper we 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.

Supporting information

S1 File. Surprisal-based analysis.
(PDF)

S2 File. 1-gram frequency analysis.
(PDF)

Author Contributions

Conceptualization: RW SF AvdB.

Data curation: AL RW.

Formal analysis: AL.

Funding acquisition: RW AvdB.

Investigation: RW.

Methodology: AL RW SF AvdB.

Software: AL RW AvdB.

Supervision: RW SF AvdB.

Visualization: AL.

Writing – original draft: AL.

Writing – review & editing: AL RW SF AvdB.

References

1. Jackendoff R. Foundations of language: brain, meaning, grammar, evolution. Oxford University Press; 2002. <https://doi.org/10.1093/acprof:oso/9780198270126.001.0001> PMID: 15377127
2. Jackendoff R. A Parallel Architecture perspective on language processing. *Brain Research*. 2007; 1146:2–22. <https://doi.org/10.1016/j.brainres.2006.08.111> PMID: 17045978
3. Hagoort P. MUC (Memory, Unification, Control) and beyond. *Front Psychology*. 2013; 4:416. <https://doi.org/10.3389/fpsyg.2013.00416> PMID: 23874313
4. Hagoort P. On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*. 2005; 9(9):416–423. <http://dx.doi.org/10.1016/j.tics.2005.07.004>. PMID: 16054419
5. Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, et al. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*. 2006; 30(4):1414–1432. <http://dx.doi.org/10.1016/j.neuroimage.2005.11.002>. PMID: 16413796
6. Kemmerer DL. *Cognitive Neuroscience of Language*. Psychology Press. 2014. Available from: <https://books.google.nl/books?id=nYi9oAEACAAJ>.
7. Bar M. Predictions in the brain: using our past to generate a future. Oxford University Press. 2011. <https://doi.org/10.1093/acprof:oso/9780195395518.001.0001>
8. Bubic A, von Cramon DY, Schubotz RI. Prediction, cognition and the brain. *Frontiers in Human Neuroscience*. 2010; <https://doi.org/10.3389/fnhum.2010.00025> PMID: 20631856
9. Friston K, Kiebel S. Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2009; 364(1521):1211–1221. <https://doi.org/10.1098/rstb.2008.0300> PMID: 19528002
10. Summerfield C, Egnér T. Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*. 2009; 13(9):403–409. <https://doi.org/10.1016/j.tics.2009.06.003> PMID: 19716752
11. Clark A. Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. *Behavioral and Brain Sciences*. 2013; 36(3):181–204. <https://doi.org/10.1017/S0140525X12000477> PMID: 23663408
12. Piantadosi ST, Tily H, Gibson E. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*. 2011; 108(9):3526–3529. <https://doi.org/10.1073/pnas.1012551108> PMID: 21278332
13. Mahowald K, Fedorenko E, Piantadosi ST, Gibson E. Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*. 2013; 126(2):313–318. <https://doi.org/10.1016/j.cognition.2012.09.010> PMID: 23116925
14. Hale J. 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. 2001;
15. Levy R. Expectation-based syntactic comprehension. *Cognition*. 2008; 106(3):1126–1177. <https://doi.org/10.1016/j.cognition.2007.05.006> PMID: 17662975
16. Frank SL, Thompson RL. Early effects of word surprisal on pupil size during reading. In: *Cognitive Science Society*; 2012. p. 1554–1559.
17. Monsalve IF, Frank SL, Vigliocco G. 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. EACL'12. Stroudsburg, PA, USA: Association for Computational Linguistics*; 2012. p. 398–408. Available from: <http://dl.acm.org/citation.cfm?id=2380816.2380866>.
18. Frank SL. Uncertainty Reduction as a Measure of Cognitive Load in Sentence Comprehension. *Topics in Cognitive Science*. 2013;(5):475–494. <http://dx.doi.org/10.1111/tops.12025>. PMID: 23681508
19. Smith NJ, Levy R. The effect of word predictability on reading time is logarithmic. *Cognition*. 2013; 128(3):302–319. <https://doi.org/10.1016/j.cognition.2013.02.013> PMID: 23747651
20. Boston MF, Hale J, Kliegl R, Patil U, Vasishth S. Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*. 2008; 2(1).
21. Frank SL, Bod R. Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*. 2011; 22:829–834. <https://doi.org/10.1177/0956797611409589> PMID: 21586764
22. Frank SL, Otten LJ, Galli G, Vigliocco G. The ERP response to the amount of information conveyed by words in sentences. *Brain Lang*. 2015; 140:1–11. <https://doi.org/10.1016/j.bandl.2014.10.006> PMID: 25461915

23. Parviz M, Johnson M, Johnson B, Brock J. Using Language Models and Latent Semantic Analysis to Characterise the N400m Neural Response; 2011.
24. Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*. 2000; 4(12):463–470. [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6) PMID: 11115760
25. Willems RM, Frank SL, Nijhof AD, Hagoort P, van Den Bosch A. Prediction During Natural Language Comprehension. *Cerebral Cortex*. 2016; 26(6):2506–2516. <https://doi.org/10.1093/cercor/bhv075> PMID: 25903464
26. Nastase S, Iacovella V, Hasson U. Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Human Brain Mapping*. 2013; 35(4):1111–1128. <https://doi.org/10.1002/hbm.22238> PMID: 23408389
27. Oostdijk N. 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); 2000. Available from: <http://www.lrec-conf.org/proceedings/lrec2000/pdf/110.pdf>.
28. Poser BA, Koopmans PJ, Witzel T, Wald LL, Barth M. Three dimensional echo-planar imaging at 7 Tesla. *NeuroImage*. 2010; 51(1):261–266. <http://dx.doi.org/10.1016/j.neuroimage.2010.01.108>. PMID: 20139009
29. Haeseryn W, Romijn K, Geerts G, de Rooij J, van den Toorn M. *Algemene Nederlandse Spraakkunst*. ONijhoff and Deurne: Wolters Plantyn; 1997.
30. Schäfer R, Bildhauer F. Building Large Corpora from the Web Using a New Efficient Tool Chain. In: Calzolari N, Choukri K, Declerck T, Dogan MU, Maegaard B, Mariani J, et al., editors. LREC. European Language Resources Association (ELRA); 2012:486–493. Available from: <http://dblp.uni-trier.de/db/conf/lrec/lrec2012.html#SchaeferB12>.
31. Stolcke A. SRILM—An Extensible Language Modeling Toolkit; 2002:901–904.
32. Daelemans W, van Den Bosch A. Memory-Based Learning in Natural Language Processing. *Memory-Based Language Processing*. 3–14. <https://doi.org/10.1017/cbo9780511486579.002>
33. Busser B, Daelemans W, van den Bosch A. Machine learning of word pronunciation: the case against abstraction. In: Sixth European Conference on Speech Communication and Technology, EURO-SPEECH 1999, Budapest, Hungary, September 5–9, 1999; 1999. Available from: http://www.isca-speech.org/archive/eurospeech_1999/e99_2123.html.
34. Baayen HR, Piepenbrock R, Gulikers L. The CELEX Lexical Database. Release 2 (CD-ROM). Philadelphia, Pennsylvania: Linguistic Data Consortium, University of Pennsylvania; 1995.
35. Friston K. Analysis of fMRI Time-Series Revisited. *NeuroImage*. 1995; 2(1):45–53. <https://doi.org/10.1006/nimg.1995.1007> PMID: 9343589
36. Keuleers E, Brysbaert M, New B. SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*. 2010; 42(3):643–650. <https://doi.org/10.3758/BRM.42.3.643> PMID: 20805586
37. Bennett CM, Wolford GL, Miller MB. The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience*. 2009; 4(4):417. <https://doi.org/10.1093/scan/nsp053> PMID: 20042432
38. Woo CW, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*. 2014; 91:412–419. <http://dx.doi.org/10.1016/j.neuroimage.2013.12.058>. PMID: 24412399
39. Nichols T, Brett M, Andersson J, Wager T, Poline JB. Valid conjunction inference with the minimum statistic. *Neuroimage*. 2005; 25:653–660. <https://doi.org/10.1016/j.neuroimage.2004.12.005> PMID: 15808966
40. Mendoza JE. In: Kreutzer JS, DeLuca J, Caplan B, editors. *Heschl's Gyrus*. New York, NY: Springer New York; 2011. p. 1243–1243. Available from: http://dx.doi.org/10.1007/978-0-387-79948-3_737.
41. Alario FX, Chainay H, Lehericy S, Cohen L. The role of the supplementary motor area (SMA) in word production. *Brain Research*. 2006; Mar 3 (1076(1)):129–143. <https://doi.org/10.1016/j.brainres.2005.11.104> PMID: 16480694
42. Willems RM, Hagoort P. Neural evidence for the interplay between language, gesture, and action: a review. *Brain and language*. 2007; 101(3):278–289. <https://doi.org/10.1016/j.bandl.2007.03.004> PMID: 17416411
43. Hertrich I, Dietrich S, Hermann A. The role of the supplementary motor area for speech and language processing. *Neurosci Biobehav Rev*. 2016; Sep (68):602–610. <https://doi.org/10.1016/j.neubiorev.2016.06.030> PMID: 27343998

44. Tremblay P, Small SL. On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *NeuroImage*. 2011; 57(4):1561–1571. <http://dx.doi.org/10.1016/j.neuroimage.2011.05.067>. PMID: 21664275
45. Tremblay P, Deschamps I, Baroni M, Hasson U. Neural sensitivity to syllable frequency and mutual information in speech perception and production. *NeuroImage*. 2016; 136:106–121. <http://dx.doi.org/10.1016/j.neuroimage.2016.05.018>. PMID: 27184201
46. Sato M, Tremblay P, Gracco VL. A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*. 2009; 111(1):1–7. <http://dx.doi.org/10.1016/j.bandl.2009.03.002>. PMID: 19362734
47. Dronkers NF, Wilkins DP, Van Valin RDJ, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. *Cognition*. 2004; 92(1–2):145–177. <http://dx.doi.org/10.1016/j.cognition.2003.11.002>. PMID: 15037129
48. Indefrey P, Cutler A. In: *Prelexical and lexical processing in listening*; 2004.
49. Snijders TM, Vosse T, Kempen G, van Berkum JJA, Petersson KM, Hagoort P. Retrieval and Unification of Syntactic Structure in Sentence Comprehension: an fMRI Study Using Word-Category Ambiguity. *Cerebral Cortex*. 2008; 19(7):1493–1503. <https://doi.org/10.1093/cercor/bhn187> PMID: 19001084
50. Pallier C, Devauchelle AD, Dehaene S. Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*. 2011; 108(6):2522–2527. <https://doi.org/10.1073/pnas.1018711108> PMID: 21224415
51. Obleser J, Meyer L, Friederici AD. Dynamic assignment of neural resources in auditory comprehension of complex sentences. *NeuroImage*. 2011; 56(4):2310–2320. <https://doi.org/10.1016/j.neuroimage.2011.03.035> PMID: 21421059
52. Brennan J, Nir Y, Hasson U, Malach R, Heeger DJ, Pykkänen L. Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*. 2012; 120(2):163–173. <https://doi.org/10.1016/j.bandl.2010.04.002> PMID: 20472279
53. Cohen L. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*. 2002; 125(5):1054–1069. <https://doi.org/10.1093/brain/awf094> PMID: 11960895
54. Price C, Devlin J. The myth of the visual word form area. *Neuroimage*. 2003; 19(3):473–481. [https://doi.org/10.1016/S1053-8119\(03\)00084-3](https://doi.org/10.1016/S1053-8119(03)00084-3) PMID: 12880781
55. Binder JR, Desai RH, Graves WW, Conant LL. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*. 2009; 19(12):2767–2796. <https://doi.org/10.1093/cercor/bhp055> PMID: 19329570
56. Binder JR, Desai RH. The neurobiology of semantic memory. *Trends in Cognitive Sciences*. 2011; 15(11):527–536. <https://doi.org/10.1016/j.tics.2011.10.001> PMID: 22001867
57. Anderson AJ, Bruni E, Bordignon U, Poesio M, Baroni M. 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*. Seattle, Washington, USA: Association for Computational Linguistics; 2013:1960–1970. Available from: <http://www.aclweb.org/anthology/D13-1202>.
58. Anderson AJ, Bruni E, Lopopolo A, Poesio M, Baroni M. Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *NeuroImage*. 2015; 120:309–322. <https://doi.org/10.1016/j.neuroimage.2015.06.093> PMID: 26188260
59. Thompson-Schill S, Ramscar M, Chrysikou E. Cognition Without Control: When a Little Frontal Lobe Goes a Long Way. *Current Directions in Psychological Science*. 2009; 18(5):259–263. <https://doi.org/10.1111/j.1467-8721.2009.01648.x> PMID: 20401341
60. Grodzinsky Y, Santi A. The battle for Broca's region. *Trends in Cognitive Sciences*. 2008; 12(12):474–480. <http://dx.doi.org/10.1016/j.tics.2008.09.001>. PMID: 18930695
61. Bornkessel-Schlesewsky I, Schlesewsky M. Linguistic Sequence Processing and the Prefrontal Cortex. *The Open Medical Imaging Journal*. 2012; 6(1-M2):47–61. <https://doi.org/10.2174/1874347101206010047>
62. Brennan JR, Stabler EP, Van Wagenen SE, Luh WM, Hale JT. Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*. 2016; 157–158:81–94. <http://dx.doi.org/10.1016/j.bandl.2016.04.008>. PMID: 27208858
63. Friederici AD, Rüschemeyer SA, Hahne A, Fiebach CJ. The Role of Left Inferior Frontal and Superior Temporal Cortex in Sentence Comprehension: Localizing Syntactic and Semantic Processes. *Cerebral Cortex*. 2003; 13(2):170. <https://doi.org/10.1093/cercor/13.2.170> PMID: 12507948
64. Herrmann B, Obleser J, Kalberlah C, Haynes JD, Friederici AD. Dissociable neural imprints of perception and grammar in auditory functional imaging. *Human Brain Mapping*. 2012; 33(3):584–595. <https://doi.org/10.1002/hbm.21235> PMID: 21391281