Research report

Rapid syntactic pre-activation in Broca’s area: Concurrent electrophysiological and haemodynamic recordings

Pelle Söderström a,b,*Merle Horne a, Peter Mannfolk c, Danielle van Westen d, Mikael Roll a

a Department of Linguistics, Centre for Languages and Literature, Lund University, Box 201, 221 00 Lund, Sweden
b MARCS Institute, Western Sydney University, Sydney, Australia
c Skane University Hospital, Department of Medical Imaging and Physiology, Lund, Sweden
d Lund University, Skane University Hospital, Department of Clinical Sciences, Diagnostic Radiology, Lund, Sweden

Article info

Article history:
Received 4 December 2017
Received in revised form 4 May 2018
Accepted 4 June 2018
Available online 5 June 2018

Keywords:
ERP
fMRI
Prediction
Speech processing
Prosody
Syntax

Abstract

Listeners are constantly trying to predict what the speaker will say next. We concurrently measured the electrophysiological and haemodynamic correlates of syntactic pre-activation, investigating when and where the brain processes speech melody cues to upcoming word order structure. Pre-activation of syntactic structure was reflected in a left-lateralised pre-activation negativity (PrAN), which was subserved by Broca’s area in the left inferior frontal gyrus, as well as the contiguous left anterior insula.

1. Introduction

Listeners constantly make predictions when processing incoming speech (Bar, 2007; DeLong et al., 2005; Dikker and Pykkänen, 2013; Federmeier, 2007; Fruchter et al., 2015; Staub and Clifton, 2006; Van Berkum et al., 2005; Wicha et al., 2004). A number of studies have investigated the neuroanatomical substrates underlying language pre-activation, strongly suggesting roles for left inferior frontal gyrus and insula as important areas involved in syntactic predictions (Henderson et al., 2016; Jakuszeit et al., 2013; Matchin et al., 2017). However, while electrophysiological studies have found that syntactic processing is supported by rapid structural predictions and that rapid analysis could be made possible by pre-parsing mechanisms of some kind (Lau et al., 2006), fewer have focused on real-time effects associated with on-going predictive activity in the brain.

1.1. Pre-activation negativity (PrAN)

We recently isolated a brain potential – the pre-activation negativity (PrAN) – elicited by word stem tones that constitute strong predictive cues to upcoming grammatical suffixes at the word level (Söderström et al., 2017a). In line with proposals by Engel et al. (2001) regarding evidence for predictive brain processes, PrAN has been suggested to reflect some aspect(s) of predictive processing – e.g. that it is sensitive to a reduction of entropy leading to fewer upcoming choices (Hale, 2006) – since its amplitude correlates with a number of both “on-going” and subsequent behavioural events (such as decision certainty regarding upcoming items, as well as response time and response accuracy (Roll et al., 2015; Söderström et al., 2016; Söderström et al., 2017b)). In addition, in Swedish – a language in which there is a strong connection between word stem tones and upcoming suffixes – the effect has been found to increase as a linear function of the number of possible upcoming word continuations given a predictive word stem tone (Söderström et al., 2016).

1.2. Clause-initial tones

In the present study, we investigated prosodic phenomena at the sentence level to determine whether clause-initial tones (Myrberg, 2010; Roll, 2006) – assumed to be important cues to syntactic structure – could be seen to pre-activate brain areas associated with grammatical processing. Previous investigations have found that more syntax-constraining or informative sentence-level tones give rise to event-related potential (ERP)
negativities (Roll and Horne, 2011; Roll et al., 2009). Building upon this, in the present study, a concurrent ERP/functional magnetic resonance imaging (fMRI) paradigm was used to investigate the neural underpinnings of syntactic pre-activation with high temporal (ERP) and spatial (fMRI) resolution.

Clause-initial tones (‘left-edge boundary tones’) in Swedish function as cues to upcoming syntactic structure (Roll et al., 2009). This is particularly important for the processing of embedded clauses starting with ‘att’ ‘that,’ where both subordinate clause (SC) and main clause (MC) structure are possible. Specifically, a low left-edge tone signals subordinate clause (SC) structure, while a high tone cues main clause structure (MC). Word order is an important cue to clause type. In SC, adverbials such as the negation inte (‘not’) precede the verb (ADV-VERB): Han hävdar att Caesar inte intog Gallien (lit. ‘He claims that Caesar not conquered Gaul’), whereas in MC, the negation follows the verb (VERB-ADV): Han hävdar att Caesar intog inte intog inte tåget (lit. ‘He claims that Caesar conquered not not the train’). These sentences are identical up to the syllable -sar in Caesar, which is either pronounced with a low clause-initial tone (signalling SC structure) or a high tone (signalling MC structure). Thus, a unique feature of left-edge boundary tones for the study of pre-activation mechanisms is that, although semantically void in themselves, they function as predictive cues to syntactic structure, i.e. upcoming word order. The low tone is more constraining since it can only pre-activate one clause type – subordinate clause – which has a fixed word order (SUBJ-ADV-VERB). The high tone, on the other hand, cues a larger class of main clause structures, including subject-initial assertive clauses (SUBJ-VERB-ADV) and verb-initial questions (VERB-SUBJ-ADV), as well as object-initial and adverb-initial topicalised clauses.

In Roll et al. (2009), a study which used test sentences similar to those presented above, the more constraining tonal cue (low left-edge boundary tone) was found to elicit an anterior ERP negativity at 200–250 ms after tone onset. While the ERP effect in that study had originally been viewed as a positivity for high tones, it has since been re-interpreted as a negativity for the more syntactically constraining low tones, following results obtained in e.g. Söderström et al. (2016). In addition to the boundary tone effect, acceptability judgements have shown that invalidly cued main clause word order (‘VERB-ADV’) after low boundary tones is less acceptable than subordinate clause word order (‘ADV-VERB’) following high boundary tones (Roll et al., 2011). This result is compatible with ideas in information theory which assume that highly weighted information (e.g. a low tone in this case) reduces entropy, whereas information with lower weight (a high tone in this context) leaves more possible outcomes (Shannon and Weaver, 1949). Invalidly cued structures after low tones also gave rise to P600 effects, indicating syntactic re-analysis, possibly triggered by the syntactic prediction made possible by the more constraining low left-edge boundary tone. The fact that a P600 was only found for word orders invalidly cued by the low tone in Roll et al. (2011) provides further support for the hypothesis that previous results can be explained by low left-edge boundary tones being more constraining as compared to high tones at the beginning of clauses. This is also in line with the previous boundary tone findings in Swedish which support the view of the P600 as reflecting re-analysis as a result of a disconfirmed prediction, rather than being a marker of grammaticality (Knoeferle et al., 2008).

1.3. Present study

The hypothesis for the present study was that low tones at the left edge of embedded clauses give rise to increased anticipation of a specific upcoming clause structure, i.e. a syntactic PrAN effect for subordinate clause structure (ADV-VERB). We tested an implication of this: the low clause-initial tone should yield increased activity in brain regions associated with syntactic processing, such as left prefrontal areas (Goucha and Friederici, 2015; Henderson et al., 2016; Jakuszeit et al., 2013; Matchin et al., 2017). Another test implication of the hypothesis is a P600 effect in the ERPs (Roll et al., 2009; Van Petten and Luka, 2012) when the expected structure is not encountered. Specifically, if, during parsing, the low left-edge boundary tone creates a stronger prediction as regards the upcoming word order in embedded clauses, we would expect to see a larger P600 effect for invalidly cued word orders (‘VERB-ADV’) after low boundary tones compared to high tones, as was found in Roll et al. (2011). Participants listened to sentences similar to those presented above, but with a pseudoword (such as kakytet) carrying the critical left-edge boundary tone in order to remove any possible semantic cues to word order. An example sentence is Rut hävdar att kakytet inte intog/intog inte tåget (‘Rut claims that the kakyt not occupied/occupied not the train’). There were four conditions in a 2 x 2 design (please refer to Fig. 1 for an illustration of the experimental design): LoValid (low boundary tone followed by validly cued SC structure), LoInvalid (low boundary tone followed by invalidly cued MC structure), HiValid (high boundary tone followed by validly cued MC structure) and HiInvalid (high boundary tone followed by invalidly cued SC structure). Participants were asked to judge as quickly as possible which word order they heard (inte intog or intog inte).

2. Results

2.1. Behavioural results

The factors used in the response time analysis were Tone (low/high left-edge boundary tone) and Validity (validly/invalidly cued word order). Neither the effect of Validity (F(1,15) = 3.385, p = 0.078, etasq = 0.193) nor the Tone x Validity interaction were significant (F(1,15) = 3.088, p = 0.099, etasq = 0.171), but as an exploratory step to probe the direction of the behavioural results, follow-up tests were performed. These showed a Validity effect on low tone stimuli (F(1,15) = 5.944, p = 0.028, etasq = 0.284), LoValid M = 969 ms (SD = 260), LoInvalid M = 1018 ms (SD = 228), suggesting that the word order violation led to longer response times following the low tone. However, caution should be taken in interpreting this effect, since the Tone x Validity interaction was not significant. No effect was found for Validity in high tone conditions (F(1,15) = 0.632, p = 0.439, HiValid M = 999 ms (SD = 239), HiInvalid M = 982 ms (SD = 260)). Thus, results suggest that LoInvalid stimuli were processed more slowly compared to LoValid. Further, no response time differences were found for word order violations in the less predictive condition (high tone).

2.2. ERP results

In the boundary tone ERP analysis, a Tone x Laterality interaction was found (F(1,15) = 11.760, p = 0.004, etasq = 0.439). Low
boundary tones gave rise to a pre-activation negativity (PrAN) over left-lateralised electrodes at 136–280 ms in the ERPs compared to high tones (F(1,15) = 7.252, p = 0.017) (Fig. 2). No effect was found over right-lateralised electrodes. This corresponded to increased neural activity seen in global root mean squares (gRMS) peaks at 100–150 ms (F(1,15) = 5.691, p = 0.031) and 150–230 ms (F(1,15) = 5.264, p = 0.037). At 450–750 ms after the syntactic disambiguation point, a Tone × Validity × Laterality interaction was found (F(1,15) = 5.590, p = 0.032, etasq = 0.272). Interaction over left-lateralised electrodes (F(1,15) = 8.678, p = 0.010, etasq = 0.366) between Tone and Validity showed that invalidly cued word order following a low boundary tone produced a P600 over left electrodes (F(1,15) = 6.682, p = 0.021). No effect was found over right-lateralised electrodes.

2.3. fMRI results

In the fMRI, the low boundary tone – in contrast to the high tone – elicited increased BOLD response primarily in a brain region associated with syntactic processing (see Table 1). The largest cluster spanned the left inferior frontal gyrus (Brodmann area (BA) 44) and the connected left anterior insula (Fig. 2). Subject variability in the gRMS amplitude of the second peak downstroke was also found to correlate positively with BOLD activity in the left frontal cluster (r = 0.609, p = 0.048, Fig. 4), coinciding in time with the second negative-going deflection, leading to the peak difference in the ERP data. As in Roll et al. (2015), the lack of upstroke correlations could indicate that the gRMS peak is related to the onset of stimulus-evoked neural activity (Skrandies, 1990). No further significant correlations were found. Invalidly cued syntactic structures showed activity in areas including the left prefrontal cortex (BA47) as well (Table 2), possibly indicating increased syntactic processing load (Fig. 3).

![Figure 2](image1)

**Fig. 2.** Top left: Scalp topography for the syntactic PrAN effect. Top right: BOLD activity for the low-high tone contrast. Bottom left: ERPs showing a PrAN beginning at 136 ms following predictive tone onset. Bottom right: gRMS amplitude was greater for low boundary tones in a 100–230 ms time window.

![Figure 3](image2)

**Fig. 3.** Top left: Scalp topography for the P600 found for main clause word orders invalidly cued by low boundary tones. Bottom left: P600 effect in the ERPs. Right: BOLD activity for the invalid-valid contrast.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Clusters found for low boundary tones (cluster extent threshold was 50 voxels).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voxels</td>
<td>MAX (z-score)</td>
</tr>
<tr>
<td>489</td>
<td>4.8</td>
</tr>
<tr>
<td>85</td>
<td>4.19</td>
</tr>
<tr>
<td>55</td>
<td>4.24</td>
</tr>
<tr>
<td>50</td>
<td>3.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Clusters found for invalidly cued word orders (cluster extent threshold was 50 voxels).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voxels</td>
<td>MAX (z-score)</td>
</tr>
<tr>
<td>111</td>
<td>4.14</td>
</tr>
<tr>
<td>71</td>
<td>3.82</td>
</tr>
<tr>
<td>65</td>
<td>4.62</td>
</tr>
</tbody>
</table>
3. Discussion

Low boundary tones (more predictively constraining) gave rise to an early left frontal negativity (PrAN) in the ERPs, as well as activity in the left BA44 and anterior insula. We thus suggest that PrAN elicited by tonal information in clause-initial words indexes pre-activation of the more strongly predicted subordinate clause (ADV-VERB) structure. The presence of a P600 in the more structurally constraining mismatch condition provides further support for the hypothesis that the low tone pre-activates syntactic structure more strongly, since a mismatch between the pre-activated and encountered structures leads the brain to re-analyse, reparse and update the analysis (Roll et al., 2011; Van Petten and Luka, 2012). Left BA44 is the core part of syntax-related Broca's area (Goucha and Friederici, 2015), and the insula has previously also been implicated in syntactic processing, especially with regard to predictive processes (Henderson et al., 2016; Jakuszeit et al., 2013; Matchin et al., 2017; Moro et al., 2001; Rodd et al., 2015).

In an investigation using a syntactic violation paradigm, Jakuszeit et al. (2013) proposed that the left inferior frontal cortex is a key area for generating top-down predictions during the course of syntactic analysis, something which possibly involves prediction-related changes in neuronal activation levels (Engel et al., 2001).

The large prefrontal and insular activations in the present study also tie in well with the strongly left-lateralised PrAN topography, as well as the ERP topography found for the contrast low minus high tone (Fig. 4). Standardised (z-score) plot showing the positive correlation in subject activity as well as the ERP topography found for the low boundary tone in a cluster.

4. Experimental procedure

4.1. Materials and methods

4.1.1. Stimuli and task

The stimuli were recorded in an anechoic chamber by a male speaker of Central Swedish and consisted of sentences with embedded clauses that were either subordinate clauses (with ADV-VERB word order) or main clauses (with VERB-ADV word order), such as Rut hållde att kakytet inte intog inte tåget ('Rut claims that the kakyt not occupied/occupied not the train').
To ensure capturing the syntactic pre-activation signal free from any semantic information, the left-edge boundary tone was carried by the final syllable of a pseudoword subject (such as kakyet, ‘the kakyt’), which began and ended with a voiceless plosive in order to facilitate cross-splicing. Four different conditions were created by cross-splicing: LoValid (low tone, SC structure), LoLinald (low tone, MC structure), HiValid (high tone, MC structure) and HiLinald (high tone, SC structure). There were 40 items in each condition for a total of 160 sentences. The negation inte and past tense verb intog were used in every sentence, with 40 different pseudowords carrying the critical tones. Sentences were identical up to the onset of the tone-bearing syllable (-tet in the example given above). Thus, the boundary tone was the first cue to upcoming clause structure. On average, the low tone involved a fall from 5.7 semitones (st, SD = 0.9) to 2.3 st (SD = 0.8) while the high tone involved a fall from 7.4 st (SD = 0.9) to 4.3 st (SD = 0.7) in the critical syllable (e.g. -tet). Mean critical syllable duration was 145 ms (SD = 14) for both low and high tone syllables. Mean sentence duration up to the critical tone-bearing syllable was 1197 ms (SD = 69) in all conditions. Mean duration of in- was 224 ms (SD = 12) in conditions LoValid and LoLinald and 196 ms (SD = 10) in HiValid and HiLinald. Nineteen native speakers of Central Swedish (mean age 24.5 years, SD = 3.9, 11 female) listened to sentences and were asked to judge as quickly as possible whether the word order in their interpretation of the embedded clause was inte intog (SC: ADV-VERB) or intog inte (MC: VERB-ADV) by pressing the left or right button on a response box. Button-response association was counter-balanced within subjects. Response times were measured from the syntactic disambiguation point –te or -tog and analysed using a within-subjects repeated-measures analysis of variance (ANOVA). Two time points were used for ERP and fMRI measurements. The first point was boundary tone onset (500 ms ERP epoch), and the second point (800 ms ERP epoch) was the same disambiguation point as for response times (burst of tec or -toc).

4.1.2. Electroencephalography (EEG)

An fMRI-compatible 32-electrode EEG system (Brain Products GmbH) was used. Sampling rate was 5 kHz, later resampled to 500 Hz before analysis. Impedances were kept below 5 kΩ. A 0.01 Hz high-pass filter was applied online and a 30 Hz low-pass filter was applied offline. Due to technical issues, EEG data from 3 participants had to be excluded, leaving 16 participants for further EEG and response time analysis (mean age 24.1 years, SD = 4.0, 10 female). Regions of interest (ROIs) used in the ERP analysis corresponded to the topographical factors laterality (left, right) and anteriority (anterior, central, posterior): left anterior (F7 and FC5), right anterior (F8 and FC6), left central (T7 and C3), right central (T8 and C4), left posterior (P7 and CP5) and right posterior (P8 and CP6). The time window for analysis of the proposed pre-activation negativity expected for low boundary tones was the same as in Roll et al. (2015), 136–280 ms after tone onset. The factors used in the analysis were Tone (low, high) and Validity (valid, invalid), as well as topographical factors (Antpost (anterior, central, posterior) and Laterality (left, right)). All and only significant effects were reported. Greenhouse-Geiser correction was used where applicable. An analysis with fewer factors – e.g. omitting the topographical factors – would have reduced the risk of Type I errors (Luck and Gasperlin, 2017). However, we believe that the analysis is motivated for the present purpose of replicating ERP results previously obtained for the same tonal difference (Roll and Horne, 2011; Roll et al., 2009, 2011) and relating them to brain areas localised using the BOLD contrast. A time window of 450–750 ms after the word order disambiguation point onset was used for the P600 analysis (following visual inspection, and in line with Roll et al. (2011)). In order to further investigate whether the low or the high tone was the main driver of neural activity, a global root mean squares (gRMS) analysis (within-subjects ANOVA) was carried out on the three peaks present in the data (Skrandies, 1990) in the time windows 50–100 ms, 100–150 ms and 150–230 ms (encompassing the respective upstroke-peak-downstroke sequences found in the data). Furthermore, the relationship between gRMS peak data and BOLD (maximum percent signal change per subject) in the largest significant cluster for the contrast low-high boundary tone was interrogated using correlation analyses of subject variability in four 20 ms time windows in gRMS: peak upstroke (beginning of peak time window) and downstroke (end of time window) focusing on the two peaks found in the significant gRMS time window. BOLD and ERP both measure time points relative to neural activity following stimulation. Therefore, reliable, concurrent variation in both haemodynamic and neurophysiological effects at the individual level for the exact same stimulus difference giving rise to stable correlations between the two types of data could stem from individual variation in the level of engagement of the same brain sources (Goense and Logothetis, 2008; Horovitz et al., 2002; Logothetis et al., 2001; Murta et al., 2015). There are also other possible sources of correlation. Thus, the ERPs might show a rapidly fading effect early in the perceptual stream, where increased signal also leads to greater, more sustained activity later in the stream. The BOLD signal would in this case be more likely to reflect the brain sources of the later effect, although it might also correlate with individual variation in the early effect. We correlated the BOLD effect with gRMS, since this signal has been argued to show the onset of “synchronous (or almost synchronous) activity of a large neuronal population in a given area” (Lehmann and Skrandies, 1980), activity which might likely be reflected in the BOLD response as well. In order to investigate the timing of the effect more closely, the two gRMS peaks were divided into upstroke and downstroke portions of 20 ms each, as in Roll et al. (2015). Since peaks in gRMS data have been suggested to reflect event-related changes in neuronal activity (Skrandies, 1990), correlations between downstroke portions of peaks and BOLD could be used to argue that the peak itself reflects the onset rather than offset of the observed effect. With regard to the BOLD data, maximum percent signal change was used, as it has previously been found to be more reliable as compared to mean percent signal change (Friedman et al., 2008). To control for the four comparisons (first peak upstroke/downstroke, second peak upstroke/downstroke), Bonferroni correction was applied, and corrected p-values are reported.

4.1.3. Functional magnetic resonance imaging (fMRI)

An event-related fMRI design was used, with trigger points for tone onset and subsequent syntactic disambiguation point identical to those used for ERP measurements, and separate fMRI analyses for the respective time-locking points (tone onset and disambiguation point). Pseudo-randomised stimulus presentation lists were created using OptSeq2 (Dale, 1999), improving deconvolution of the haemodynamic response. Stimulus-onset asynchrony between stimulus sentences was jittered between 4 and 8 s. Structural and functional MRI data was acquired using a Siemens Prisma 3 T scanner and a 64-channel head coil (structural image: T1 MPRAGE (TR = 1900 ms, TE = 2.5 ms, flip angle = 9°, matrix size = 256 × 256, field of view = 256 × 256 mm², 176 slices, slice thickness = 1 mm) and functional data: T2* gradient-echo EPI (TR = 2100 ms, TE = 30 ms, flip angle = 75°, matrix size = 96 × 96, field of view = 192 × 192 mm², 25 slices, slice thickness = 4 mm)). Data was analysed using FSL 5.0.9 (www.fmrib.ox.ac.uk/fsl). High-pass filtering (128 s), motion correction (MCFLIRT (Jenkinson et al., 2002)), slice-timing correction and 6 mm FWHM Gaussian smoothing was applied to the data. A gamma function was used to model the haemodynamic response function. The temporal derivative was included as a within-subject regressor at the
Acknowledgements

We are grateful to the editor and to two anonymous reviewers for insightful comments and suggestions for changes which have led to considerable improvements in the manuscript.

This work was supported by Knut and Alice Wallenberg Foundation (Grant No. 2014.0139), Marcus and Amalia Wallenberg Foundation (Grant No. 2014.0039), and the Swedish Research Council (Grants No. 2011–2284 and 2017–06384).

References


This work was supported by Knut and Alice Wallenberg Foundation (Grant No. 2014.0139), Marcus and Amalia Wallenberg Foundation (Grant No. 2014.0039), and the Swedish Research Council (Grants No. 2011–2284 and 2017–06384).
