Title

How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition

Authors

Tanner, Darren
Morgan-Short, Kara
Luck, Steven J

Peer reviewed

How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition

Darren Tanner,
Department of Linguistics, Beckman Institute for Advanced Science and Technology, Neuroscience Program, University of Illinois at Urbana-Champaign

Kara Morgan-Short, and
Department of Hispanic and Italian Studies, Department of Psychology, University of Illinois at Chicago

Steven J. Luck
Department of Psychology, Center for Mind and Brain, University of California Davis

Abstract

Although it is widely known that high-pass filters can reduce the amplitude of slow ERP components, these filters can also introduce artifactual peaks that lead to incorrect conclusions. To demonstrate this and provide evidence about optimal filter settings, we recorded ERPs in a typical language processing paradigm involving syntactic and semantic violations. Unfiltered results showed standard N400 and P600 effects in the semantic and syntactic violation conditions, respectively. However, high-pass filters with cutoffs at 0.3 Hz and above produced artifactual effects of opposite polarity before the true effect. That is, excessive high-pass filtering introduced a significant N400 effect preceding the P600 in the syntactic condition, and a significant P2 effect preceding the N400 in the semantic condition. Thus, inappropriate use of high-pass filters can lead to false conclusions about which components are influenced by a given manipulation. The present results also lead to practical recommendations for high-pass filter settings that maximize statistical power while minimizing filtering artifacts.

The progress of science requires that experimental results are valid, replicable, and generalizable. In any domain of experimental inquiry, however, decisions must be made about data cleaning and processing procedures, and the choices made in any given study could potentially impact whether these goals are realized. In the domain of human electrophysiology, researchers using electroencephalography (EEG) and event-related brain potentials (ERPs) to study cognition must make numerous choices about recording and post

Corresponding Author: Darren Tanner, Department of Linguistics, 4080 Foreign Languages Building, MC-168, 707 S. Mathews Ave, Urbana, IL 61810, dstanner@gmail.com.

processing parameters, including when and how to apply filters to their data.[1] Although all filters distort time-domain data to some extent, filtering can be beneficial by removing frequency components unlikely to be of cortical origin (e.g., muscle activity, skin potentials, or electrical line noise) thereby improving signal-to-noise in the data and statistical power (Kappenman & Luck, 2010). Additionally, some filtering of high frequencies is necessary during digitization in order to avoid aliasing artifacts (see Luck, 2014). In other words, the benefits of filtering outweigh the costs when appropriate filter parameters are used.

However, to meet the goals of producing valid findings that accurately reflect the brain processes being recorded, it is important for researchers to understand how some filter settings may lead to significant distortion of the ERP waveforms and thereby result in misleading conclusions. Whereas some recent discussion has focused on low-pass filters (i.e., filters that attenuate high frequencies) and how they may impact conclusions about the time course of ERP effects (e.g., Rousselet, 2012; VanRullen, 2011; Widmann & Schröger, 2012), our goal here is to assess the impact of high-pass filters (i.e., filters that attenuate low frequencies) on slower cortical potentials such as the N2, the P300, the N400, the late positive potential, etc (cf. Acunzo, MacKenzie, & van Rossum, 2012; Widmann, Schröger, & Maess, in press). For the sake of concreteness, we will focus our analyses on ERP components that are commonly studied in language comprehension (the N400 and P600), demonstrating that misapplication of high-pass filters can distort the ERP waveform in ways that lead to spurious conclusions about the engagement of the cognitive processes under investigation. However, our conclusions apply broadly to all slow components.

Since the seminal paper of Duncan-Johnson and Donchin (1979), it has been widely known that excessive use of high-pass filters can reduce the amplitude of late, slow endogenous components. They showed that as the high-pass filter cutoff frequency became progressively higher, the amplitude of the P300 elicited by deviant tone bursts became correspondingly smaller. This amplitude reduction was apparent already with time constants as long as 1 second (∼0.16 Hz high-pass cutoff). More recently, Hajcak and colleagues (Hajcak, Weinberg, MacNamara, & Foti, 2012) provided a similar caution about amplitude reductions caused by high-pass filters in the context of the late positive potential (LPP) elicited by emotionally arousing stimuli. In their study, LPP amplitude was markedly attenuated with a .5 Hz high-pass filter and virtually eliminated with a 1 Hz filter. While these effects on ERP amplitudes are of clear methodological significance, they also have larger implications for theory building, in that cross-study comparisons and meta-analysis of ERP effect magnitudes become difficult when filter cutoffs vary across reported experiments. Although filtering requirements will necessarily differ depending on factors such as the ERP component under investigation, the recording environment, and the participant population, it is important for researchers to have a clear understanding of how their filter settings might distort their waveforms and lead to incorrect or misleading conclusions.

---

[1]The term "filter" can be applied to many types of signal processing procedures. We use the term in this paper to refer to the most common class of filters used in EEG/ERP processing, which can be described in terms of a frequency response function that specifies the amount of attenuation for each frequency band.

Although it is widely known that high-pass filters may attenuate ERP effects, it is less widely known that these filters may induce artifactual effects into the ERP waveform. In other words, high-pass filters may not merely lead to the absence of an effect, they may lead to the presence of a statistically significant but artifactual effect. More specifically, high-pass filters can produce artifactual deflections of opposite polarity and/or artifactual oscillatory activity before and after an elicited experimental effect of interest.

Figure 1 shows examples of these filter distortions with an artificial ERP waveform consisting of a single cycle of a cosine wave, which resembles a P600 waveform (see below).[2] When this waveform is low-pass filtered with a half-amplitude cutoff of 2 Hz (to remove high frequencies), this effectively "spreads out" the voltage in time, causing the onset to start earlier and the offset to end later (Figure 1A). Panel B depicts the result of applying a high-pass filter with the same half-amplitude cutoff at 2 Hz. The filters depicted in Panels A and B are exactly opposite; one passes low frequencies and the other attenuates low frequencies. Consequently, the distortion produced by the high-pass filter has the opposite polarity of the distortion produced by the low-pass filter(see Luck, 2014, for a detailed discussion of these filtering artifacts; see also Widmann et al., in press, for a more technical overview of digital filter implementation). In other words, whereas the spreading of the waveform produced by the low-pass filter in Panel A is positive in polarity (because the waveform being filtered consists of a single positive peak), the spreading of the waveform produced by the high-pass filter in Panel B is negative.

The "inverted spreading" produced by high-pass filters is particularly problematic, because it introduces new peaks into the waveform, and this may lead researchers to conclude that a given experimental manipulation is impacting multiple components rather than a single component. Thus, high-pass filtering does not merely reduce the amplitude of a slow component, it can create artifactual peaks that lead to completely invalid conclusions. Note that filters with steeper roll-offs may produce oscillating artifacts and not just single peaks (Bénar, Chauvière, Bartolomei, & Wendling, 2010; Yeung, Bogacz, Holroyd, Nieuwenhuis, & Cohen, 2007).

It is important to note that the magnitude of these filter distortions depends on the cutoff frequencies of the filters, with greater distortion as more of the signal is attenuated by the filter. For example, Figures 1C and 1D show that almost no visible distortion is produced in this P600-like waveform by a 30 Hz low-pass filter or a 0.01 Hz high-pass filter. Thus, high-pass filtering with reasonably low filter cutoffs may be a useful procedure that attenuates low-frequency noise without introducing meaningful distortion of the waveforms.

To show how the distortion depends on the specific cutoff frequency, Figure 2 overlays the effects of six different filter settings (0.01 Hz, 0.1 Hz, 0.3 Hz, 0.5 Hz., 0.7 Hz, and 1 Hz; 24 dB/octave roll-off) on the same simulated P600 effect. There were virtually no differences between the 0.01 and 0.1 Hz filters (and these waveforms were nearly identical to the unfiltered waveform; see Figure 1). However, beginning with the 0.3 Hz filters and above,

---

[2]Note that this waveform is similar in properties to other slower cortical potentials, such as the P300, LPP, etc.

the peak amplitude decreased approximately linearly as the cutoff increased, and this was accompanied by artifactual negative peaks on either side of the positive peak.

The negative peak preceding the simulated P600 wave has a latency that is very similar to the latency of the N400 component (see below). Consequently, the use of a high-pass filter cutoff of 0.3 Hz or higher may create an artifactual peak that is interpreted as being an N400. Moreover, the artifactual N400-like peak causes the apparent onset of the true P600 peak to increase in latency, and this latency shift increases as the high-pass filter cutoff increases, with a 1 Hz filter delaying the onset of the positivity by approximately 200 ms. Similarly, the application of such a filter to a P300 wave could lead to an artifactual peak at 200 ms that is interpreted as an N2 wave, and this would also increase the apparent onset latency of the P300. Thus, even a relatively 'mild' 0.3 Hz filter can produce artifactual peaks and alter the time course of a real effect onset.

Not all variation in high-pass filters necessarily leads to excessive distortions. High-pass cutoff frequencies of 0.1 Hz and lower have relatively little effect on ERP components, including slower endogenous components like the P300 and LPP (Acunzo et al., 2012; Hajcak et al., 2012; Kappenman & Luck, 2010; Widmann et al., in press). Indeed a widely cited methodological review for the elicitation of commonly-studied slower ERP components—including the mismatch negativity (MMN), P300, and N400—recommends data recording parameters with high-pass filter settings between 0.01 and 0.1 Hz (Duncan et al., 2009), and similar recommendations are made in a widely-used introductory text on ERPs (Luck, 2005, 2014). Some recent studies and methodological reviews in the clinical and cognitive EEG literature have also cautioned against the use of excessive high-pass filters precisely because of the possibility for introducing artifacts (Acunzo et al., 2012; Bénar et al., 2010; Yeung et al., 2007; see also Widmann et al., in press). Whereas the vast majority of studies report filter cutoffs within the recommended range of 0.01 – 0.1 Hz, even a brief survey of papers published in highly regarded journals, including respected neuroscience and electrophysiology journals, will turn up numerous ERP papers with filters exceeding these values, including cutoff values as high as 2 Hz despite these cautions (e.g., Nenonen, Shestakova, Huotilainen, & Näätänen, 2005). Part of the problem may be that artifactual effects induced by filtering may not readily be recognized as artifactual, as they could resemble theoretically relevant and/or expected ERP effects. Empirical data are needed to demonstrate that these artifacts have a practical impact. The goal of the present study is therefore to make the problems associated with excessive high-pass filtering more concrete by showing how they may lead to inappropriate conclusions about cognitive processing in real data from a typical language ERP paradigm. We will also provide evidence about filter settings that optimize statistical power without producing significant distortion.

Although the conclusions of the present study are not specific to language research, it is necessary to provide some background about language-related ERP components so that the implications of the present results are clear. Much language-related ERP research has focused on characterizing two ERP components: N400 and P600. This dichotomy is important, because the N400 and P600 effects have traditionally been associated with different levels of linguistic processing. Modulations of the N400 are traditionally associated

with processing at the semantic level, where amplitude is typically inversely related to the ease of accessing or integrating words and their semantic features in a given context (e.g., Holcomb, 1993; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980, 1984). The N400 is often contrasted with the P600, which has most frequently been associated with difficulty in processing grammar or morphosyntax (e.g., Hagoort, Brown, & Osterhout, 1999; Osterhout & Holcomb, 1992), and more recently, syntactically well-formed but propositionally impossible sentences (e, g, Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Kuperberg, 2007; Paczynski & Kuperberg, 2012; van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; van Herten, Chwilla, & Kolk, 2006). In some reports, lexical/semantic and grammatical manipulations have had independent and additive effects on the N400 and P600, respectively, indicating that these two components reflect—at least to a first approximation—separable and independent neurocognitive processes (Allen, Badecker, & Osterhout, 2003; Osterhout & Nicol, 1999).

Because of this property, many studies have used the N400/P600 dichotomy as a criterion for establishing whether a given linguistic manipulation engages processing mechanisms that are primarily semantic in nature (producing an N400 effect) or primarily syntactic in nature (producing a P600 effect) (e.g., Hagoort, Wassenaar, & Brown, 2003; Kim & Osterhout, 2005; Osterhout & Mobley, 1995). Other studies have used these two effects as indices of stages of neurocognitive development for language processing in language learners(e.g., Morgan-Short, Steinhauer, Sanz, & Ullman, 2012; Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013). Thus, theoretical inferences often crucially rely on determining whether a given manipulation impacts primarily the N400, primarily the P600, or both components.

The artifactual, opposite-polarity peaks that can be produced by high-pass filters are particularly problematic for experiments that require distinguishing between N400 and P600 effects. As shown in Figures 1 and 2, high-pass filters can take a late positive effect such as the P600 and introduce an artifactual negative component at an earlier latency that might be mistaken for an N400. In other words, a purely syntactic P600 effect might be distorted to produce an artifactual negative peak after filtering, which in turn might lead to an incorrect conclusion about the presence of an N400 effect in the waveform. Similarly, a purely semantic N400 effect might be preceded by an artifactual positive peak that would be interpreted as an unusually early P2 effect, or it might be followed by an artifactual positive peak that would be interpreted as a P600. This potential for confusion is especially likely given that the N400 and P600 exhibit a large degree of overlap in their scalp topographies.

As mentioned above, reported filter settings in ERP studies of language comprehension (as well as in other cognitive domains) vary widely. Although most filter cutoffs are at or below 0.1 Hz, numerous reported studies from prominent labs in high profile journals use filter settings ranging from .25 Hz to .5 Hz, with some reports using cutoffs as high as 1 or 2 Hz, with no clear justification provided for these relatively high cutoff values. One possible justification is that higher cutoffs might more effectively remove skin potentials and therefore increase statistical power (Kappenman & Luck, 2010; see Loerts, Stowe, & Schmid, 2013, for an example of this sort of justification). However, it is necessary to

balance the improvement in statistical power with the possibility of introducing artifactual effects.

For these reasons, it is therefore necessary to empirically determine exactly how various high-pass filter settings can both lead to spurious effects and improve statistical power for late, slow endogenous ERP components, such as the N400 and P600. To accomplish this, we conducted an experiment using well-studied semantic and syntactic violations to elicit N400 and P600 effects, respectively. We then applied a broad range of different high-pass filters to the resulting data to assess how these filters distorted the ERP waveforms. We also used these data to perform Monte Carlo simulations of experiments with different numbers of participants, which allowed us to determine how statistical power varied across different filter settings. Together these two sets of analyses made it possible to determine the optimal balance between statistical power and waveform distortion.

Method

Participants

Participants were 33 native English speakers. All were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision, and reported no history of neurological impairment. As described in detail below, data from nine participants were excluded from analysis because of artifacts, leaving 24 participants (mean age = 19 years, SD = 1.17; 11 female) in the final analysis. All participants provided informed consent and received course credit for taking part.

Materials

Experimental materials were English sentences that were well-formed, or contained a subject-verb agreement or lexical-semantic violation (presented in an unpredictable order). For subject-verb agreement sentences, 70 sentence pairs were created. One version of each pair was well-formed, and the other contained a violation of subject-verb agreement (e.g., *Many doctors claim/*claims that insurance rates are too high*). All sentences had plural subjects, such that ungrammatical verbs were in singular form, and there were no cues to sentence grammaticality prior to the onset of the critical verb.[3] Sentence pairs were counterbalanced across two lists in a Latin square design, such that each list had 35 well-formed and 35 ill-formed items, and no single participant saw the same version of each sentence.

For the semantic condition, 70 sentences were created with passive structures (e.g., *The rough part of the wood was sanded to perfection*). Violation sentences were created by exchanging the verb in one sentence with the verb in another (e.g., *The unpleasant cough syrup was swallowed while the boy held his nose*), such that the verb was no longer semantically coherent with the preceding context (e.g., *The rough part of the wood was # swallowed to perfection/The unpleasant cough syrup was # sanded while the boy held his*

---

[3]We use the * and # symbols to indicate stimuli that are ungrammatical and semantically/pragmatically anomalous, respectively. Critical words for ERP averages are underlined here to make the critical ERP word apparent to the reader, but were presented without underlines or */# symbols during the experiment.

nose). Sentence pairs were counterbalanced across two lists, such that there were 35 semantically well-formed and 35 ill-formed sentences per list. Each participant saw each critical verb twice, once in a well-formed and once in an ill-formed sentence. However, each sentence frame was only presented to each participant once.

Each list of sentences contained 140 total sentences. Both lists were pseudorandomized such that no more than three well-formed or ill-formed sentences followed in immediate succession, and no more than two sentences from the same condition followed in succession. The two lists were presented either in forward or reverse order, to avoid list order effects across participants.

Procedure

Participants were tested individually in a single session. Each participant was randomly assigned to one of the stimulus lists and seated in a comfortable chair at a viewing distance of 120 cm from a CRT monitor. Participants were instructed to read each sentence silently while remaining relaxed and minimizing movements. Each trial began with "Ready?" displayed in the center of the monitor, and participants pressed a button on a mouse to begin the trial. The trial sequence consisted of a fixation cross, followed by each word of the sentence presented one at a time in rapid serial visual presentation format. The fixation cross and each word were presented in the center of the screen, and remained on the screen for 350 ms with a 150 ms blank screen between words. Words were presented in Arial font (approximately .5° of visual angle high and no more than 3°wide). Sentence-ending words were presented with a period, followed by a "Good/Bad" screen, instructing participants to give an acceptability judgment for the sentence via mouse button press. Participants were instructed to respond "good" if they felt the sentence was well-formed in English or "bad" if they felt the sentence was in any way anomalous (e.g., grammatically or semantically odd). After each response, the "Ready?" screen appeared again, allowing participants to begin the next trial at their own pace. Experimental trials were preceded by ten practice trials.

EEG recording and analysis—Continuous EEG was recorded from 30 Ag/AgCl electrodes embedded in an elastic cap (ANT Waveguard caps) from standard and extended 10-20 locations (Jasper, 1958: FP1, FPz, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, POz, O1, Oz, O2), as well as from electrodes placed on the left (M1) and right (M2) mastoids. EEG was recorded on-line to a common average reference, and re-referenced off-line to the averaged activity over M1 and M2. Vertical eye movements and blinks were recorded from a bipolar montage consisting of electrodes placed above and below the left eye, and horizontal eye movements were recorded from a bipolar montage consisting of electrodes placed at the outer can thus of each eye. Impedances at each electrode site were held below 5 kΩ. The EEG was amplified with an ANT Neuro bioamplifier system (AMP-TRF40AB Refa-8 amplifier), digitized with a 512 Hz sampling rate and filtered with a digital finite impulse response low-pass filter with a 138.24 Hz cutoff (sampling rate * .27), as implemented by ASA-lab recording software (v4.7.9; ANT Neuro Systems). No high-pass filter was applied during data acquisition (i.e., recordings were direct coupled).

All offline data processing was carried out using the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) Matlab toolboxes. Filtering was accomplished via ERPLAB's Butterworth filters, which are infinite impulse response filters that are applied bidirectionally to achieve zero phase shift. Fourth order filters were used to achieve an intermediate roll-off slope of 24 dB/octave (after the second pass of the filter). The filters were applied to the continuous EEG to avoid edge artifacts. The data were low-pass filtered at 30 Hz in all cases, and with seven different high-pass cutoffs: DC (no filter), 0.01 Hz, 0.1 Hz, 0.3 Hz, 0.5 Hz, 0.7 Hz, and 1 Hz. Note that the half-amplitude (-6 dB) point is used to describe all filter cutoffs in the present paper.

ERPs were computed off-line for each participant in each condition, time-locked to the onset of the critical verb in each sentence (underlined, above), relative to a 200 ms prestimulus baseline. Trials characterized by excessive ocular, motor, or other artifact were excluded from analysis. Artifact detection was carried out on epochs in two stages beginning 200 ms prior to and ending 1000 ms following the onset of the critical word. In stage one, a moving window peak-to-peak threshold criterion was applied to the two bipolar eye channels. The window was 200 ms wide and was advanced in 100 ms increments. Epochs with peak-to-peak amplitudes exceeding 50μV were marked for exclusion. This screening targeted large, rapid ocular artifacts, such as blinks and eye saccades. Second, scalp channels were screened for large drifts or deviations (e.g., large skin potentials, excessive motor artifact, excessive alpha); specifically, epochs in which the voltage exceeded an absolute threshold of ±100μV were excluded from analysis. So that an identical set of trials was included for each high-pass filter setting, artifact rejection was carried out on the DC-30 Hz filtered data for each participant, and then applied to each of that participant's subsequent filtered datasets. Data from any participant with more than 25% rejected trials overall were excluded from analysis (see Luck, 2014), resulting in the exclusion of nine participants from the final analyses (see above). Approximately 5.6% of trials were excluded in the 24 remaining participants.

Because the N400 and P600 effects are most prominent over centro-parietal scalp sites, we focused our ERP analyses on a centro-parietal region of interest (ROI) consisting of eight electrodes, where the N400 and P600 effects are typically largest (e.g., Tanner & Van Hell, 2014): C3, Cz, C4, CP1, CP2, P3, Pz, and P4. The waveforms from these sites were averaged prior to mean amplitude measurement. As syntactic and semantic manipulations have been shown to impact P2, N400 and P600 effects (see e.g., Osterhout & Mobley, 1995; Tanner, Nicol, & Brehm, 2014), we computed mean amplitudes for each condition in three time windows: 150-300 ms, 300-500 ms, and 500-800 ms, corresponding to the three components, respectively. These windows were chosen a priori as canonical time windows for the P2, N400, and P600 components. To simplify our analytical approach, our primary analyses focused on experimental effects by computing mean amplitudes from difference waves in the ill-formed minus well-formed sentences, separately for the semantic and syntactic conditions. Difference wave data were analyzed with one-sample t-tests for each time window and each filter setting, investigating whether the mean difference differed significantly from zero (i.e., whether an experimental effect was present or not). No correction for sphericity violations was needed for these single-df analyses. We additionally conducted a set of Monte Carlo analyses to investigate the effects of high-pass filter settings

on statistical power to detect effects. Monte Carlo simulations will be described with the results, below.

Results

Figure 3 shows grand mean ERP waveforms for the well-formed and ill-formed words in both the semantic and syntactic violation conditions in combination with each filter setting. Results from the one-sample t-test analyses are presented in Tables 1 and 2. The DC waveforms show the data without any contamination from high-pass filtering and provide a reference against which the other waveforms can be compared. In the DC waveforms, brain responses to ill-formed sentences in the syntactic conditions showed a large positive-going effect beginning approximately 500 ms poststimulus (a P600 effect), and brain responses to ill-formed sentences in the semantic condition showed a negative-going effect beginning at approximately 300 ms poststimulus (an N400 effect), which also continued throughout the recording epoch. Statistical analysis confirmed these observations (Tables 1 and 2): syntactic violations led to a significant positivity in the 500-800 ms time window, and semantic violations led to a significant negativity in the 300-500 ms window. Note that the N400 effect extended beyond the canonical measurement, leading to significant effects in the later time window; extended N400 effects of this nature have been reported in some previous studies in subsets of individuals (Kos, van den Brink, & Hagoort, 2012; Tanner & Van Hell, 2014). These results confirm that our experimental paradigm elicited the same pattern of syntactic (P600) and semantic (N400) violation effects that have been observed in numerous previous ERP studies (see Swaab, Ledoux, Camblin, & Boudewyn, 2012, for a recent review).

The effects shown in Figure 3 with no high-pass filtering were very similar when a 0.01 or 0.1 Hz high-pass filter was applied, and there was no qualitative change in statistical significance with these filters (Tables 1 and 2). However, the effects changed markedly as the cutoff was increased to 0.3 Hz and above. The effects of the filters can be compared more easily in Figure 4, in which difference waves for the ill-formed minus well-formed comparison are overlaid for the different filter cutoffs. The waveforms are very similar for DC, 0.01 Hz, and 0.1 Hz, but clear reductions in the amplitude of the N400 and P600 effects in the semantic and syntactic conditions, respectively, can be seen as the filter cutoff increases above 0.1 Hz. In addition, these amplitude reductions were accompanied by effects of opposite polarity in the time window preceding the standard effect. That is, beginning with high-pass filters above 0.3 Hz, there was a significant N400-like negativity preceding the P600 in the syntactic condition (see Table 1), and with filter settings at and above 0.7 Hz there was a significant positivity in the P2 time window preceding the N400 in the semantic condition (see Table 2). Neither of these effects (the N400 in the syntactic condition and the P2 in the semantic condition) was significant with the more conservative high-pass filter settings (i.e., those settings that attenuate only the very lowest frequencies).

Although it was not statistically significant, a small N400 effect was visible in the syntactic condition even with no high-pass filtering. This effect was due to individual differences in participants' brain responses to syntactic violations, with a subset of participants primarily showing an N400 effect to subject-verb agreement violations. These sorts of individual

differences have been previously documented (Tanner & Van Hell, 2014). This raises the possibility that the significant N400 effect observed in this condition with cutoffs of 0.3 Hz and higher is not artifactual, but instead reflects a true effect that was "revealed" by the filter. Two findings argue against this possibility. First, Figure 2 shows that high-pass filters with cutoffs of 0.3 Hz and higher produce an artifactual negativity in the N400 latency range when the waveform is know to contain only a P600-like positive wave. Second, we repeated our analyses of the syntactic condition of the experiment, using only those participants who primarily showed a P600 effect (n = 16; see Tanner & Van Hell, 2014), and filtering still produced an N400 in these participants. This is illustrated in Figure 5, which shows the grand average data from these 16 participants with no high-pass filtering (DC) and with a 0.5-Hz high-pass filter. The filter led to a more negative potential in the N400 latency range for the ill-formed sentences compared to the well-formed sentences even though no sign of this effect was present in the DC wave forms. Thus, high-pass filtering truly creates artifactual effects.

Having shown that filtering can reduce true effects and induce artifactual components, we additionally investigated whether filtering alters the scalp topographies of these effects. Topographic plots are presented in Figure 6 for anomalous minus well-formed words in the syntactic (panel A) and semantic conditions (panel B) for three filter settings: 0.1, 0.5, and 1 Hz. Visual inspection shows attenuation of the true effects, as well as the induced effects at higher filter settings, which reflects our analysis above. Visual inspection also shows that the topographies of the effects differed slightly across the different filter settings. For example, the voltage from 400-500 ms in the semantic condition (Figure 6B) was broadly distributed when a 0.1 Hz cutoff but had a clear posterior focus at higher filter settings.

To assess the statistical significance of these changes in topography, we quantified mean amplitudes in the syntactic condition in the 300-500 ms and 500-800 ms time windows for data in the 0.1 and 0.5 Hz filter conditions over four lateral ROIs: left frontal (F7, F3, FC5, FC1), right frontal (F8, F4, FC6, FC2), left posterior (P7, P3, CP5, CP1), right posterior (P8, P4, CP6, CP2). We submitted these data to a repeated measures ANOVA with two levels of condition (well-formed, anomalous), two levels of anteriority (anterior, posterior), two levels of hemisphere (left, right), and two levels of filter (0.1 Hz, 0.5 Hz). Of crucial interest was whether the factor of filter setting interacted with any of the topographic factors, and whether there were further interactions between filter, topography, and condition. We report only the interactions involving filter and the topographic factors.

In the 500-800 ms time window (which corresponds to the P600 effect), the following interactions were significant: filter by anteriority ($F(1, 23) = 22.492$, MSE = 1.195, $p < .001$), filter by condition by anteriority ($F(1, 23) = 10.139$, MSE = 0.377, $p = .004$), filter by hemisphere ($F(1, 23) = 5.593$, MSE = 0.411, $p = .027$), and filter by condition by hemisphere ($F(1, 23) = 15.715$, MSE = 0.089, $p < .001$). In the 300-500 ms time window (corresponding to the N400 effect), the following interactions were found: filter by anteriority ($F(1, 23) = 28.892$, MSE = 0.567, $p < .001$), filter by condition by anteriority ($F(1, 23) = 5.118$, MSE = 0.198, $p = .033$), filter by hemisphere ($F(1, 23) = 10.846$, MSE = 0.197, $p = .003$), and filter by condition by hemisphere ($F(1, 23) = 9.617$, MSE = 0.042, $p = .005$).

Thus, there is clear evidence that filtering altered the scalp distributions of the experimental effects of interest, both real and artifactual.

Guides to ERP methods typically recommend applying high-pass filters to continuous EEG rather than to segmented EEG or averaged ERP waveforms, which avoids "edge" artifacts (Luck, 2014). However, some studies nonetheless report applying high-pass filters to either epoched EEG or averaged ERP data. We therefore further investigated whether applying high-pass filters to the continuous EEG versus individuals' averaged ERPs might differentially impact the present data. Figure 7 compares the effects of applying a 0.1 Hz high-pass filter to the continuous EEG and averaged ERPs in the syntactic condition. As can be seen, applying this high-pass filter to the averaged ERPs led both to a reduction of P600 effect size and an increase in N400 effect size, relative to this same filter applied to continuous EEG. Note that the filter was exactly the same in both cases; the only difference was whether it was applied to the continuous EEG or the segmented and averaged ERPs. This provides empirical support for the recommendation that high-pass filters be applied to continuous EEG, rather than segmented EEG or averaged ERP waveforms.

Monte Carlo Simulations to Determine Statistical Power

Although high-pass filters can create artifactual effects, they may also increase the statistical power for detecting real effects. To provide evidence about the tradeoff between these costs and benefits of filtering, we conducted a series of Monte Carlo simulations on the ERP data that make it possible to quantify statistical power. The general logic of this approach was to simulate experiments with different numbers of participants by randomly selecting subsets of the 24 participants from the actual data set and then conducting the statistical analyses on these subsets. For example, to simulate an experiment with 10 participants, we randomly selected 10 of the 24 participants (with replacement) and conducted the statistical analyses on these 10 participants. We then repeated this over and over with different random subsets of 10 participants. This allowed us to compute the probability of a significant effect in a "generic" experiment with 10 participants. The same process was then repeated with different sample sizes so that the function relating statistical power (probability of detecting a significant effect) to sample size could be computed (see Kappenman & Luck, 2010, for a similar approach). This provides a more general assessment of statistical power than examining the significance level with the full sample of 24 participants, because it takes into account the sampling error that occurs in real experiments. In our analyses, 10,000 experiments were simulated for each sample size between 5 and 20 individuals (each selecting a random subset of the full sample of 24 participants). Separate simulations were carried out for the 300-500 and 500-800 ms time windows in the syntactic condition, and for the 150-300 and 300-500 ms time windows in the semantic condition.

Simulation results for the syntactic condition are depicted in Figure 8. The top panel shows the probability of a significantly greater positivity for the ill-formed sentences (e.g., a P600) in the 500-800 ms window (which presumably reflects a "true" P600 syntactic violation effect), and the bottom panel shows the probability of a significantly greater negativity for the ill-formed sentences in the 300-500 ms window (which presumably reflects a "bogus" N400 syntactic violation effect). The greatest statistical power to detect a true P600 effect in

the 500-800 ms window occurred with a high-pass filter of 0.1 Hz. Filter settings at 0.01 and 0.3 Hz performed somewhat poorer, followed by DC recordings (no high-pass filter) and 0.5 Hz filters, with 0.7 and 1 Hz filters not even reaching 50% power to detect an effect with 20 participants. In the 300-500 ms time window, the 0.1 Hz filter setting showed the lowest likelihood of producing a significant (but artifactual) N400 effect, followed closely by the 0.01 Hz filter and the DC recording. Beginning with the 0.3 Hz filter, there was a marked increase in likelihood to detect a significant (but artifactual) N400 effect in the syntactic violation condition, with a high likelihood of detecting this artifactual effect with filters of 0.5 Hz and above.

The Monte Carlo results from the semantic condition are depicted in Figure 9. The top panel shows the probability of a significantly greater negativity for the ill-formed sentences in 300-500 ms window (which presumably reflects a "true" N400 semantic violation effect), and the bottom panel shows the probability of a significantly greater positivity for ill-formed sentences in the 150-300 ms window (which presumably reflects a "bogus" P2semantic violation effect). Here, the probability of detecting a true effect (the N400 semantic violation effect) was greatest with the lowest filter cutoffs (DC – 0.01 Hz) and declined systematically as the cutoff increased. This was particularly striking with the 0.7 and 1 Hz filters, for which the power to detect a significant effect fell below 50% in simulated experiments with small sample sizes. Note that these less-conservative filters would also presumably lead to poor statistical power with larger sample sizes in experiments with smaller effect sizes. At the same time, filters with higher cutoff frequencies led to an increase in the probability of detecting a significant (but artifactual) P2 effect prior to the N400, just as they led to significant but artifactual negative effects prior to the P600 in the syntactic violation analyses.

Discussion

This experiment provided concrete evidence about the impact of high-pass filter settings on slow ERP components, taking two components commonly studied in language comprehension(the N400 and P600) as a case in point. Consistent with the filtering artifacts observed in the simulated data shown in Figures 1 and 2, high-pass filters above 0.3 Hz induced artifactual components into the ERP waveforms from this experiment, which would potentially lead to completely false conclusions about which components were influenced by the experimental manipulations.

More specifically, with no high-pass filtering (DC recordings) and conservative high-pass filters (0.1 Hz), syntactic violations elicited a classic P600 effect, and semantic violations elicited a classic N400 effect. However, as the high-pass filter cutoff was increased to 0.3 Hz and above, the amplitude of the P600 to syntactic violations decreased markedly, and the P600 was accompanied by a significant (but artifactual) N400 effect. At filter settings above 0.7 Hz, the P600 ceased to be significant, leaving only a significant but artifactual N400 effect in the data. Thus, these filters obscured the true effect and created a bogus effect. Similarly, in the semantic violation condition, typical and highly significant N400 effects were found with a filter cutoff of 0.1 Hz or less, and the amplitude of these effects decreased markedly with filter settings above 0.3 Hz. Beginning with filter settings of 0.7 Hz and

above, the negativity in the later of our two N400 time windows ceased to be significant, and this was accompanied by a significant (but artifactual) positivity in the P2 time window. These filtering artifacts were similar to those seen our simulated data, showing that misapplication of high-pass filters do indeed have real-life consequences for ERP studies of cognition.

It is important to stress that these spurious filter-induced artifacts would substantially change the conclusions that would be drawn about the processes engaged during language comprehension. This follows from the fact that the N400 and P600 reflect separable cognitive processes that have been associated with very different aspects of language comprehension (e.g., processing of lexical and semantic features is associated with the N400, and processing of (morpho) syntactic and combinatorial information is associated with the P600). As we have demonstrated here, inappropriate use of high-pass filters would cause enormous problems for studies examining whether a given experimental manipulation impacts primarily the N400, P600, or both, or for studies requiring differentiation between the N400 and P2 effect. In such cases, researchers would come to false conclusions about which components are impacted by their experimental manipulations. Note that these results have implications for studies of cognition beyond language, as the N400 and P600 have properties similar to several other slower endogenous components commonly studied in cognitive ERP experiments (e.g., the N200, MMN, P300, LPP, etc), and are equally applicable to analysis and interpretation of data from other electromagnetic methods, such as ERFs from magneto encephalography. Moreover, with sufficiently high filter cutoffs (e.g., 1 Hz or higher), even faster and earlier ERP components such as the N170 and the mismatch negativity will be affected in this manner. Indeed, Acunzo et al. (2012) recently provided empirical evidence from face processing that filter settings greater than 0.1 Hz can introduce artifactual short-latency ERP components. High-pass filters systematically reduced a late, slow positive component, and introduced an early-latency deflection in the 50-100 ms poststimulus time window, which could be interpreted as a C1 component. Moreover, as in our data, they showed that filters can alter the scalp topographies of the measured effects, and they concluded that even relatively modest high-pass filter settings can lead to spurious conclusions about early sensory components.

Our Monte Carlo simulations additionally provide a detailed estimate of the trade-off between the noise reduction provided by high-pass filtering and the possibility of introducing artifactual components. In the syntactic condition, the optimal trade-off occurred with a 0.1 Hz filter. In this case, power to detect a true P600 effect was maximized, whereas the likelihood of detecting a significant (but bogus) N400 effect was minimized. In the semantic condition, the picture is somewhat different. The optimal trade-off appeared to be best with minimal high-pass filtering (DC or 0.01 Hz); however filters with 0.1 Hz cutoffs showed only minimal impacts to detect true N400 effects, and still a low likelihood to detect artifactual P2 effects. Kappenman and Luck (2010) performed a similar Monte Carlo simulation examining the P300 component in an oddball task, and they also found that 0.1 Hz provided the optimal tradeoff between waveform distortion and statistical power.

The non-monotonic relationship between filter setting and statistical power in the syntactic condition (and in Kappenman & Luck, 2010) and the slight discrepancy in optimal filter

settings across conditions likely stems from the fact that high-pass filters will have the most benefit the farther away from the baseline period you are. This is because baseline correction itself can greatly attenuate low frequency noise, but this attenuation declines as the waveform drifts away from the baseline voltage over time. Consequently, low frequency noise in the data is likely to have the greatest impact on the slowest potentials, like the P600 or P300. In these cases, high-pass cutoffs of 0.1 Hz are most likely to attenuate low-frequency noise, which can contaminate these slow components, while providing minimal distortion to the underlying ERP effect. Thus, 0.1 Hz is likely to be optimal for very late components. More research is needed to establish the generality of these findings and to determine the optimal cutoffs for components with earlier latencies.

Note that although Widmann et al. (in press) generally concur with our recommendation to use high-pass filter settings 0.1 Hz, they also mention that higher filter settings can potentially be used in lieu of baseline correction, and that in some cases, attenuation of effect amplitudes from high-pass filters can be a reasonable compromise in situations where filtering can provide other methodological benefits (e.g., improved source localization). The purpose of baseline correction is to remove slow voltage offsets that might otherwise add uncontrolled variance to the data and reduce statistical power. Although high-pass filters will certainly remove these nonsystematic sources of variance, the present results demonstrate that these filters produce systematic distortions of the data that can easily lead to statistically significant but bogus conclusions. A major problem with baseline correction is that any systematic differences across conditions in the prestimulus voltage can lead to bogus effects in the poststimulus period after baseline correction is performed (see Luck, 2014). However, high-pass filtering will not actually eliminate these poststimulus effects. Instead, high-pass filtering will take the prestimulus differences and spread an inverted version of them into the poststimulus period (just as they inverted and spread the N400 and P600 effects in the present study). Thus, the best solution is typically to use well-controlled experimental designs to prevent systematic differences in prestimulus voltage and then to use the prestimulus voltage for baseline correction.

Having shown that appropriate use of high-pass filtering can improve statistical power while not introducing spurious artifactual effects, we are now in a position to provide practical advice regarding the use of high-pass filters in studies of language and other cognitive domains. First, mild attenuation of slow frequencies can improve statistical power to detect true effects, while providing little or no distortion to these effects. Our Monte Carlo simulations showed that high-pass filter settings between 0.01 and 0.1 Hz were optimal for statistical power in detecting true effects while not introducing distortions, with the 0.1 Hz filter providing the best statistical power for the longest latency ERP effect (the P600). Raising the high-pass cutoff above 0.1 Hz not only decreased statistical power for detecting true effects, but also increased the likelihood of creating artifactual effects, even when the filter was applied to the continuous EEG.

However, note that there are some cases when slow potentials are not necessarily artifactual, where even modest high-pass filtering might obscure these potentially meaningful signals. Such slow potentials that develop over several seconds include the contingent negative variation (Brunia, Van Boxtel, & Böcker, 2012; Walter, Cooper, Aldridge, McCallum, &

Winter, 1964) and the contralateral delay activity (Drew & Vogel, 2008), among others. In the domain of language comprehension, some studies have shown slow potentials developing over several words in a sentence can index meaningful cognitive activity and co-vary with experimental manipulations (King & Kutas, 1995; Kutas & King, 1996; Phillips, Kazanina, & Abada, 2005; Van Petten & Kutas, 1991). For any of these slow potentials, high-pass filtering could remove this important signal from the data. Additionally, in some cases this slow drift can co-mingle with experimental effects of interest, making it difficult to know if changes in component amplitude are due to the experimental manipulation or the slow drift. In such cases high-pass comparing filtered and unfiltered waveforms may help disentangle this separate sources of activity in the EEG(see e.g., Kutas & King, 1996; Van Petten & Kutas, 1991, for examples). However, given that we have shown here that extreme filters can introduce meaningful temporal distortions and spurious peaks in ERP waveforms, one must be cautious in the interpretation of the resulting filtered waveforms. In particular, researchers should empirically demonstrate that any effects detected after the filtering procedure were non-artifactual. Moreover, if slow potentials are present in the data, researchers should try to identify why they exist, rather than reflexively treating them as noise and removing them with filters. That is, researchers can use the slow potentials as a source of information about the underlying cognitive processes (i.e., treat them as signal).

Second, high-pass filters should be applied to continuous EEG data, not epoched EEG or averaged ERPs. Typical high-pass filters involve computations based on several seconds' worth of data, and edge artifacts can occur when the time series is not sufficiently long. When the continuous EEG is used, the time series contains a large buffer both before and after a given event, providing a sufficient period of time for the filter to operate properly. These buffers are missing from typical epoched EEG and averaged ERPs. For example, in the present data, high-pass filtering of the averaged ERPs introduced artifacts of a similar nature to extreme high-pass filter settings when compared to the same filter applied to continuous EEG (Figure 7), though these edge artifacts may manifest themselves differently in other datasets.

In conclusion, the present results make it clear that filtering is not a harmless procedure that can be applied without concern to the choice of parameters. With the wrong parameters, a true effect may not be significant, and a bogus but statistically significant effect may be introduced, leading to conclusions that are completely false. However, by following some simple guidelines, filtering can improve statistical power without producing meaningful distortion of the data, leading to valid conclusions and the ability to compare findings across studies. Thus, the use of appropriate high-pass filters is critical to the contribution that ERP research can make to advancing scientific inquiry into cognition.

Acknowledgments

References

Acunzo DJ, MacKenzie G, van Rossum MCW. Systematic biases in early ERP and ERF components as a result of high-pass filtering. Journal of Neuroscience Methods. 2012; 209(1):212–218.10.1016/j.jneumeth.2012.06.011 [PubMed: 22743800]

Allen M, Badecker W, Osterhout L. Morphological analysis in sentence processing: An ERP study. Language and Cognitive Processes. 2003; 18(4):405–430.

Bénar CG, Chauvière L, Bartolomei F, Wendling F. Pitfalls of high-pass filtering for detecting epileptic oscillations: A technical note on "false" ripples. Clinical Neurophysiology. 2010; 121(3): 301–310.10.1016/j.clinph.2009.10.019 [PubMed: 19955019]

Brunia, CHM.; Van Boxtel, GJM.; Böcker, KBE. Negative slow waves as indices of anticipation: The bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. In: Luck, SJ.; Kappenman, ES., editors. The Oxford handbook of event-related potential components. Oxford: Oxford University Press; 2012. p. 189-207.

Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods. 2004; 134(1):9–21.10.1016/j.jneumeth.2003.10.009 [PubMed: 15102499]

Drew T, Vogel EK. Neural measures of individual differences in selecting and tracking multiple moving objects. The Journal of Neuroscience. 2008; 28(16):4183–4191.10.1523/JNEUROSCI.0556-08.2008 [PubMed: 18417697]

Duncan CC, Barry RJ, Connolly JF, Fischer C, Michie PT, Näätänen R, et al. Van Petten C. Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clinical Neurophysiology. 2009; 120(11):1883–1908.10.1016/j.clinph.2009.07.045 [PubMed: 19796989]

Duncan-Johnson CC, Donchin E. The Time Constant In P300 Recording. Psychophysiology. 1979; 16(1):53–56. [PubMed: 758627]

Hagoort, P.; Brown, CM.; Osterhout, L. The neurocognition of syntactic processing. In: Brown, CM.; Hagoort, P., editors. The neurocognition of language. Oxford: Oxford University Press; 1999. p. 273-316.

Hagoort P, Wassenaar M, Brown CM. Syntax-related ERP-effects in Dutch. Cognitive Brain Research. 2003; 16(1):38–50. [PubMed: 12589887]

Hajcak, G.; Weinberg, A.; MacNamara, A.; Foti, D. ERPs and the study of emotion. In: Luck, SJ.; Kappenman, ES., editors. The Oxford handbook of event-related potential components. New York: Oxford University Press; 2012. p. 441-474.

Hoeks JCJ, Stowe LA, Doedens G. Seeing words in context: The interaction of lexical and sentence level information during reading. Cognitive Brain Research. 2004; 19(1):59–73.10.1016/j.cogbrainres.2003.10.022 [PubMed: 14972359]

Holcomb PJ. Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. Psychophysiology. 1993; 30:47–61. [PubMed: 8416062]

Jasper HH. The ten–twenty system of the International Federation. Electroencephalography and Clinical Neurophysiology. 1958; 10:371–375.

Kappenman ES, Luck SJ. The effects of electrode impedance on data quality and statistical significance in ERP recordings. Psychophysiology. 2010; 47(5):888–904.10.1111/j.1469-8986.2010.01009.x [PubMed: 20374541]

Kim A, Osterhout L. The independence of combinatory semantic processing: Evidence from event-related potentials. Journal of Memory and Language. 2005; 52(2):205–225.10.1016/j.jml.2004.10.002

King JW, Kutas M. Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. Journal of Cognitive Neuroscience. 1995; 7:376–395. [PubMed: 23961867]

Kos M, van den Brink D, Hagoort P. Individual variation in the late positive complex to semantic anomalies. Frontiers in Psychology. 2012; 3(318)10.3389/fpsyg.2012.00318

Kuperberg GR. Neural mechanisms of language comprehension: Challenges to syntax. Brain Research. 2007; 1146:23–49.10.1016/j.brainres.2006.12.063 [PubMed: 17400197]

Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology. 2011; 62:621–647.10.1146/annurev.psych.093008.131123

Kutas M, Hillyard SA. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science. 1980; 207(4427):203–205. [PubMed: 7350657]

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307:161–163. [PubMed: 6690995]

Kutas, M.; King, JW. The potentials for basic sentence processing: Differentiating integrative processes. In: Inui, T.; McClelland, JL., editors. Attention and Performance XVI. Cambridge, MA: MIT Press; 1996. p. 501-546.

Loerts H, Stowe LA, Schmid MS. Predictability speeds up the re-analysis process: An ERP investigation of gender agreement and cloze probability. Journal of Neurolinguistics. 2013; 26(5):561–580.10.1016/j.jneuroling.2013.03.003

Lopez-Calderon J, Luck SJ. ERPLAB: An open-source toolbox for the analysis of event-related potentials. Frontiers in Human Neuroscience. 2014; 8(213)10.3389/fnhum.2014.00213

Luck, SJ. An introduction to the event-related potential technique. 1st. Cambridge, MA: MIT Press; 2005.

Luck, SJ. An introduction to the event-related potential technique. 2nd. Cambridge, MA: MIT Press; 2014.

Morgan-Short K, Steinhauer K, Sanz C, Ullman MT. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. Journal of Cognitive Neuroscience. 2012; 24(4):933–947. [PubMed: 21861686]

Nenonen S, Shestakova A, Huotilainen M, Näätänen R. Speech-sound duration processing in a second language is specific to phonetic categories. Brain and Language. 2005; 92(1):26–32.10.1016/j.bandl.2004.05.005 [PubMed: 15582033]

Oldfield RC. The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Osterhout L, Holcomb PJ. Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language. 1992; 31(6):785–806.

Osterhout L, McLaughlin J, Pitkänen I, Frenck-Mestre C, Molinaro N. Novice learners, longitudinal designs, and event-related potentials: A means for exploring the neurocognition of second language processing. Language Learning. 2006; 56(Suppl. 1):199–230.

Osterhout L, Mobley L. Event-related brain potentials elicited by failure to agree. Journal of Memory and Language. 1995; 34:739–773.

Osterhout L, Nicol J. On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. Language and Cognitive Processes. 1999; 14(3):283–317.

Paczynski M, Kuperberg GR. Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. Journal of Memory and Language. 2012; 67(4):426–448.10.1016/j.jml.2012.07.003 [PubMed: 23284226]

Phillips C, Kazanina N, Abada SH. ERP effects of the processing of syntactic long-distance dependencies. Cognitive Brain Research. 2005; 22:407–428.10.1016/j.cogbrainres.2004.09.012 [PubMed: 15722211]

Rousselet GA. Does filtering preclude us from studying ERP time-courses? Frontiers in Psychology. 2012; 3(131)10.3389/fpsyg.2012.00131

Swaab, TY.; Ledoux, K.; Camblin, CC.; Boudewyn, MA. Language-related ERP components. In: Luck, SJ.; Kappenman, ES., editors. The Oxford handbook of event-related potential components. Oxford: Oxford University Press; 2012. p. 397-439.

Tanner D, McLaughlin J, Herschensohn J, Osterhout L. Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. Bilingualism: Language and Cognition. 2013; 16(2):367–382.10.1017/S1366728912000302

Tanner D, Nicol J, Brehm L. The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. Journal of Memory and Language. 2014; 76:195–215.10.1016/j.jml.2014.07.003 [PubMed: 25258471]

Tanner D, Van Hell JG. ERPs reveal individual differences in morphosyntactic processing. Neuropsychologia. 2014; 56(4):289–301.10.1016/j.neuropsychologia.2014.02.002 [PubMed: 24530237]

Van de Meerendonk N, Kolk HHJ, Vissers CTWM, Chwilla DJ. Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. Journal of Cognitive Neuroscience. 2010; 22(1):67–82. [PubMed: 19199401]

Van Herten M, Chwilla DJ, Kolk HHJ. When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. Journal of Cognitive Neuroscience. 2006; 18(7):1181–1197.10.1162/jocn.2006.18.7.1181 [PubMed: 16839291]

Van Petten C, Kutas M. Influences of semantic and syntactic context on open- and closed-class words. Memory & Cognition. 1991; 19(1):95–112. [PubMed: 2017035]

VanRullen R. Four common conceptual fallacies in mapping the time course of recognition. Frontiers in Psychology. 2011; 2(365)10.3389/fpsyg.2011.00365

Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL. Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. Nature. 1964; 203:380–384. [PubMed: 14197376]

Widmann A, Schröger E. Filter effects and filter artifacts in the analysis of electrophysiological data. Frontiers in Psychology. 2012; 3(233)10.3389/fpsyg.2012.00233

Widmann A, Schröger E, Maess B. Digital filter design for electrophysiological data - a practical approach. Journal of Neuroscience Methods. n.d. 10.1016/j.jneumeth.2014.08.002

Yeung N, Bogacz R, Holroyd CB, Nieuwenhuis S, Cohen JD. Theta phase resetting and the error-related negativity. Psychophysiology. 2007; 44(1):39–49.10.1111/j.1469-8986.2006.00482.x [PubMed: 17241139]
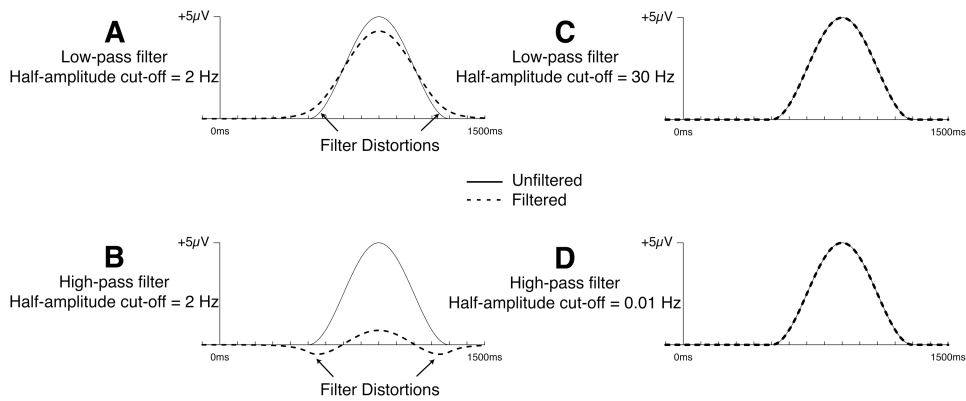
Figure 1.
Effects of filtering on a simulated P600-like ERP component. The simulated component is a single-cycle cosine wave with an amplitude of 5μV, onset of 500 ms poststimulus, and duration of 800 ms. The simulated component was embedded in 20 seconds of zero values to avoid filtering edge effects. Panels A and B show distortions caused by 2 Hz low-pass (A) and high-pass (B) filters. Panels C and D show no visible distortion to the original waveform with 30 Hz low-pass (C) and 0.01 Hz high-pass (D) filters. Filter frequencies correspond to the half-amplitude (-6 dB) cut-off (12 dB/octave roll-off).
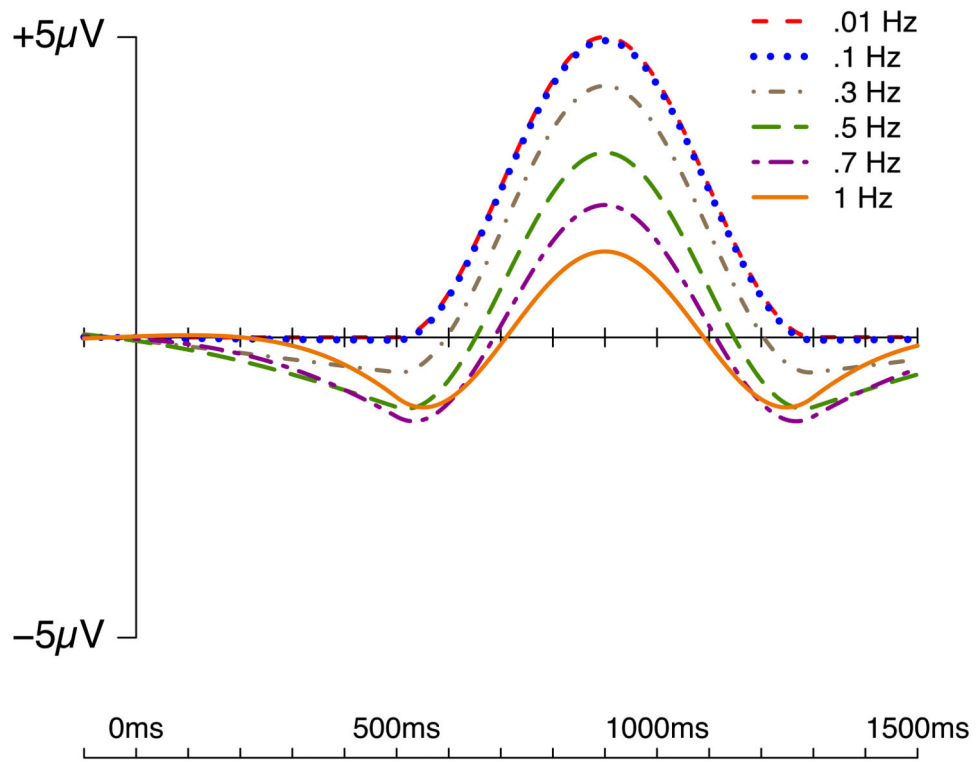
Figure 2.
Effects of different high-pass filter settings (24 dB/octave roll-off) on the simulated P600 component.
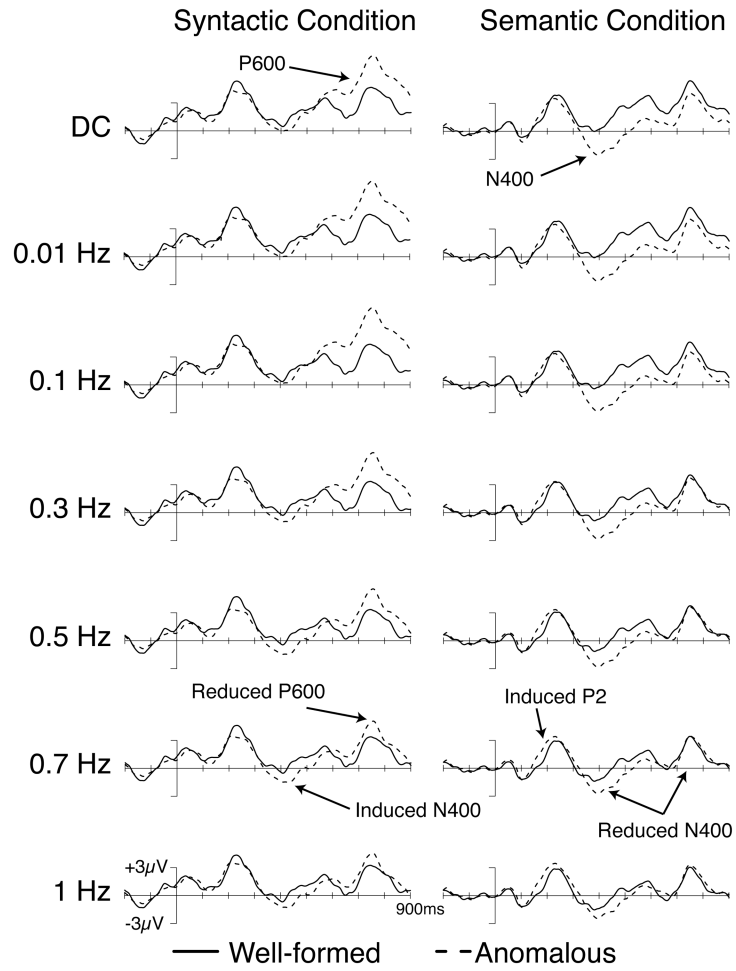
Syntactic Condition   Semantic Condition

P600

DC

N400

0.01 Hz

0.1 Hz

0.3 Hz

0.5 Hz

Reduced P600   Induced P2

0.7 Hz

Induced N400

Reduced N400

+3μV

1 Hz

-3μV   900ms

——— Well-formed   - - Anomalous

Figure 3.
Effects of high-pass filter settings on experimental data for participants' brain responses to
well-formed (solid line) and ill-formed (dashed line) verbs in the syntactic (left column) and
semantic (right column) violation conditions. Waveforms reflect the average of activity over
centro-parietal electrode sites (C3, Cz, C4, CP1, CP2, P3, Pz, P4). Onset of the verb is
indicated by the vertical bar; 900 ms of activity is depicted. Positive voltage is plotted up.
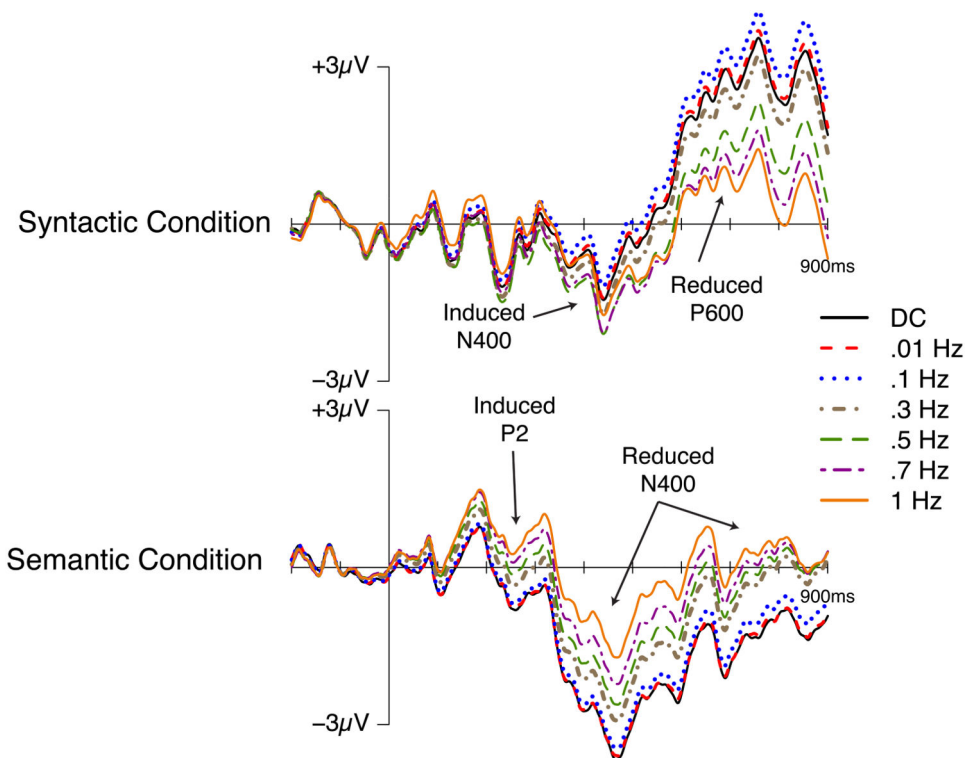
Figure 4.
Difference waves from the syntactic (upper plot) and semantic (lower plot) conditions depicting the impact of different high-pass filter settings on experimental effects. Waveforms reflect the average of activity over centro-parietal electrode sites (C3, Cz, C4, CP1, CP2, P3, Pz, P4). Onset of the verb is indicated by the vertical bar; 900 ms of activity is depicted. Positive voltage is plotted up.
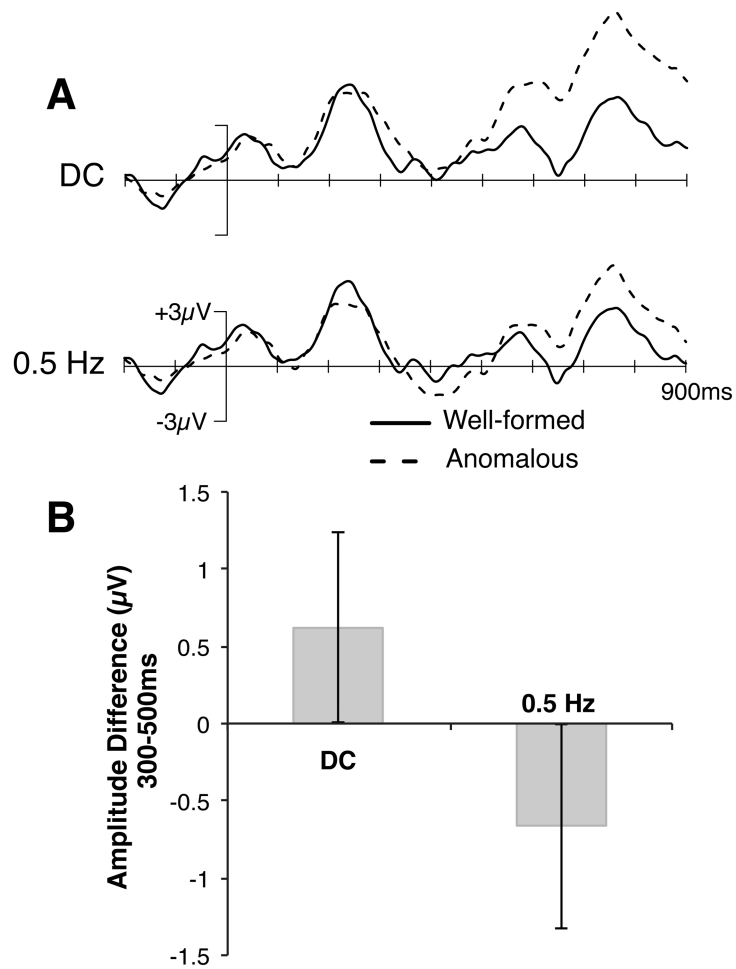
Figure 5.
Impact of high-pass filtering for only those participants showing predominantly P600 effects in the syntactic condition (n = 16). Panel A depicts waveforms with no high-pass filter (DC, upper waveform) and a 0.5 Hz high-pass filter (lower waveform). Panel B shows effect magnitudes (mean amplitude in the ungrammatical minus grammatical condition) for the N400 time window (300-500 ms). Error bars show 95% confidence intervals. ERPs showed a significant positive effect with no high-pass filter ($p < .05$) and a significant negative effect with a 0.5 Hz high-pass filter ($p < .05$) in the 300-500 ms time window.
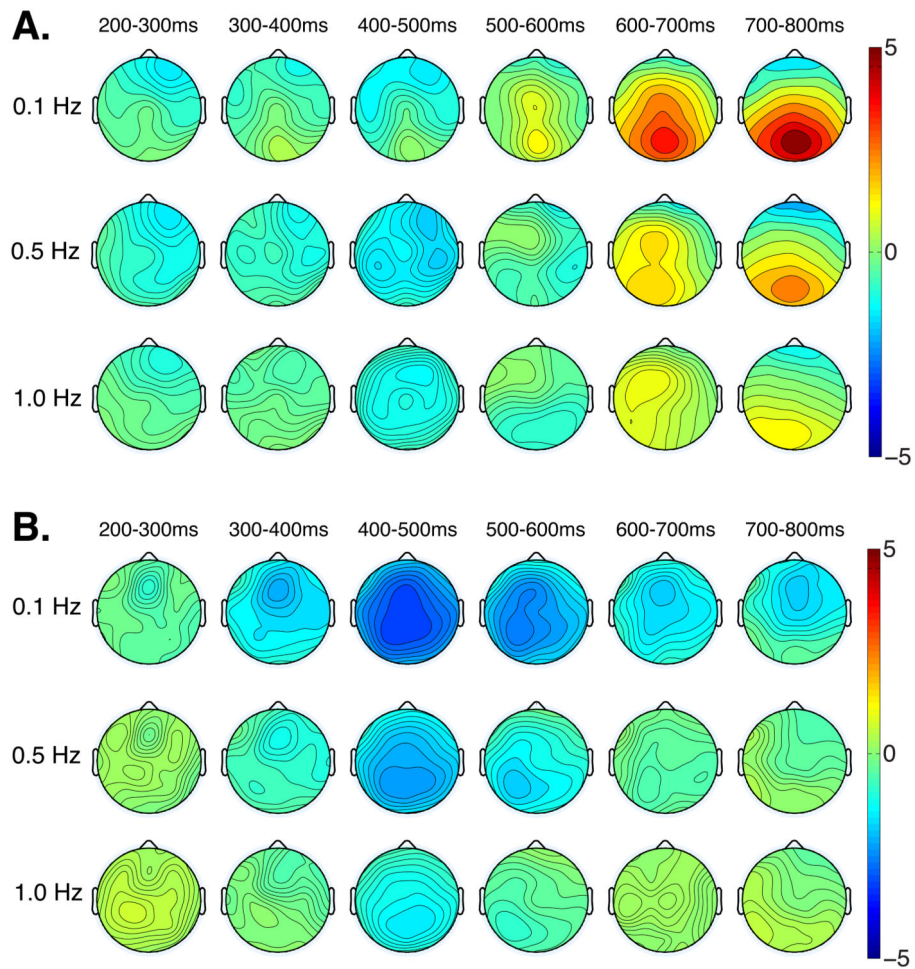
Figure 6.
Topographic distributions of experimental effects for mean amplitude in 100 ms bins for the well-formed minus ill-formed words for syntactic (A) and semantic (B) conditions for three filter settings (0.1 Hz, 0.5 Hz, and 1 Hz).
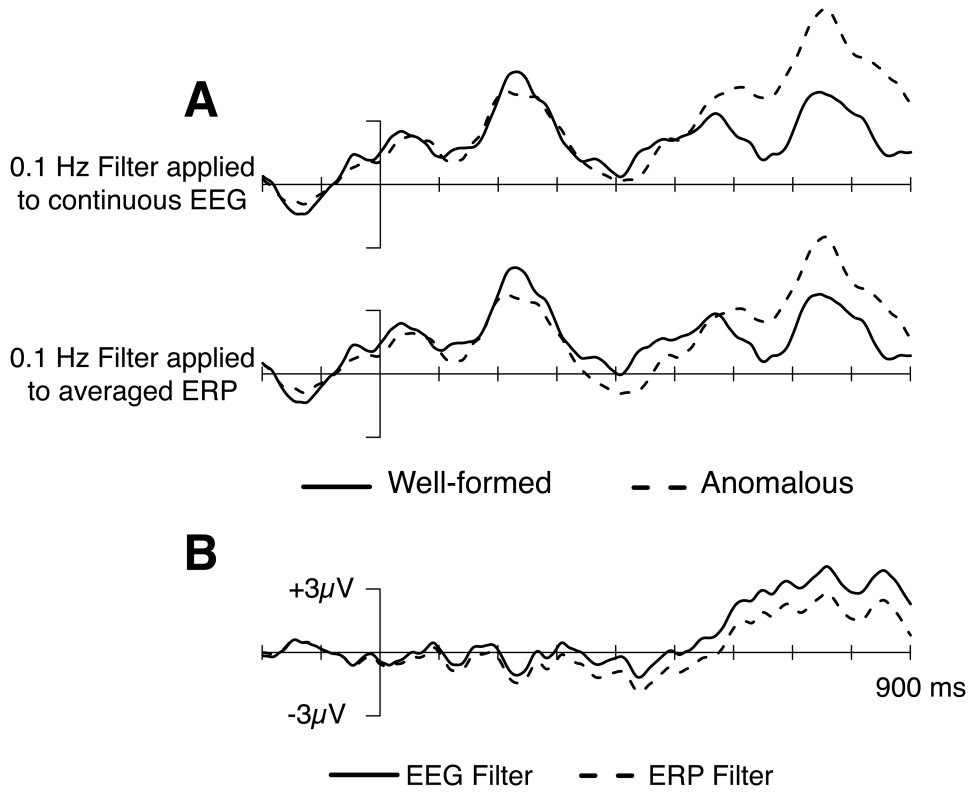
Figure 7.
Differential effect of applying a 0.1 Hz high-pass filter to the continuous EEG or averaged ERPs. Panel A depicts grand averaged ERPs for the well-formed and anomalous verbs in the syntactic condition. The upper waveform shows the 0.1 Hz filter applied to continuous EEG; the lower waveform shows the 0.1 Hz filter applied to the averaged ERPs. Panel B shows difference waves comparing effects of applying filters to the continuous EEG versus averaged ERPs.
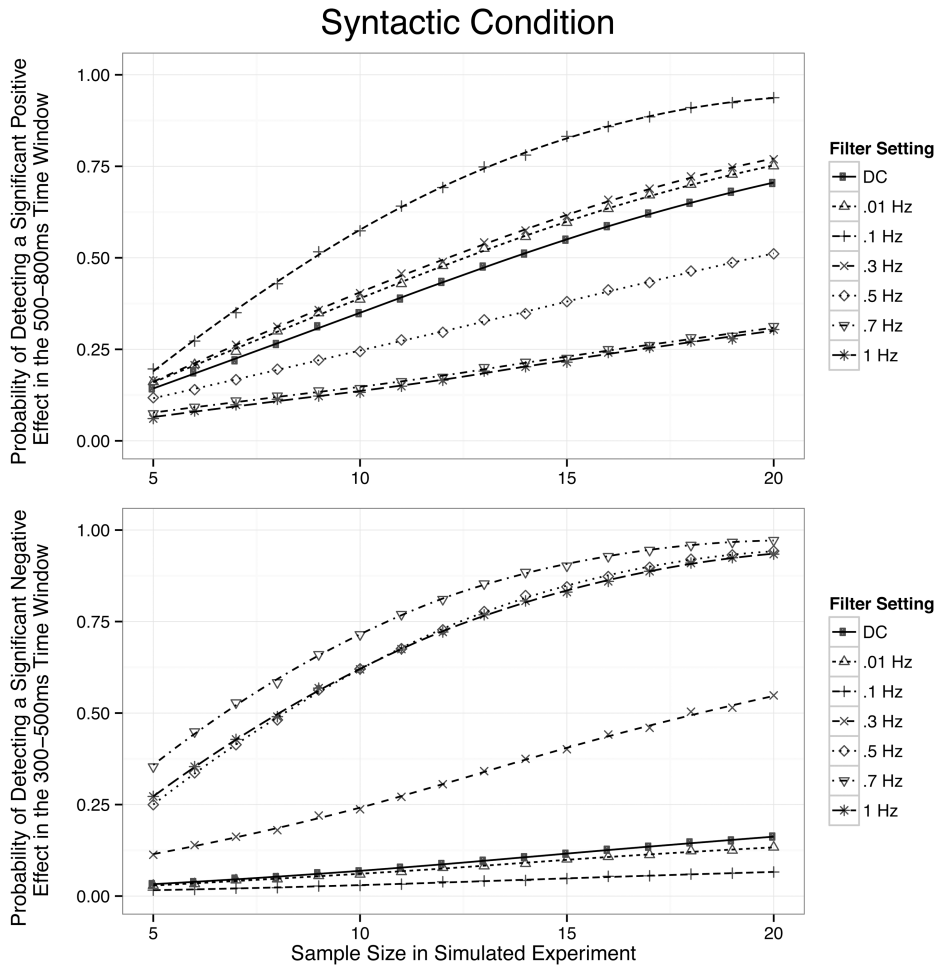
Syntactic Condition

Figure 8.
Probability of detecting significant effects in the syntactic condition as a function of sample size and filter setting, as revealed by Monte Carlo simulations across 10,000 simulated experiments between five and 20 participants. The upper panel depicts power to detect a significant positivity in the 500-800 ms time window (a P600 effect); the lower panel depicts power to detect a significant negativity in the 300-500 ms time window (an artifactual N400 effect). Smooth lines are fit by local regression (LOESS).
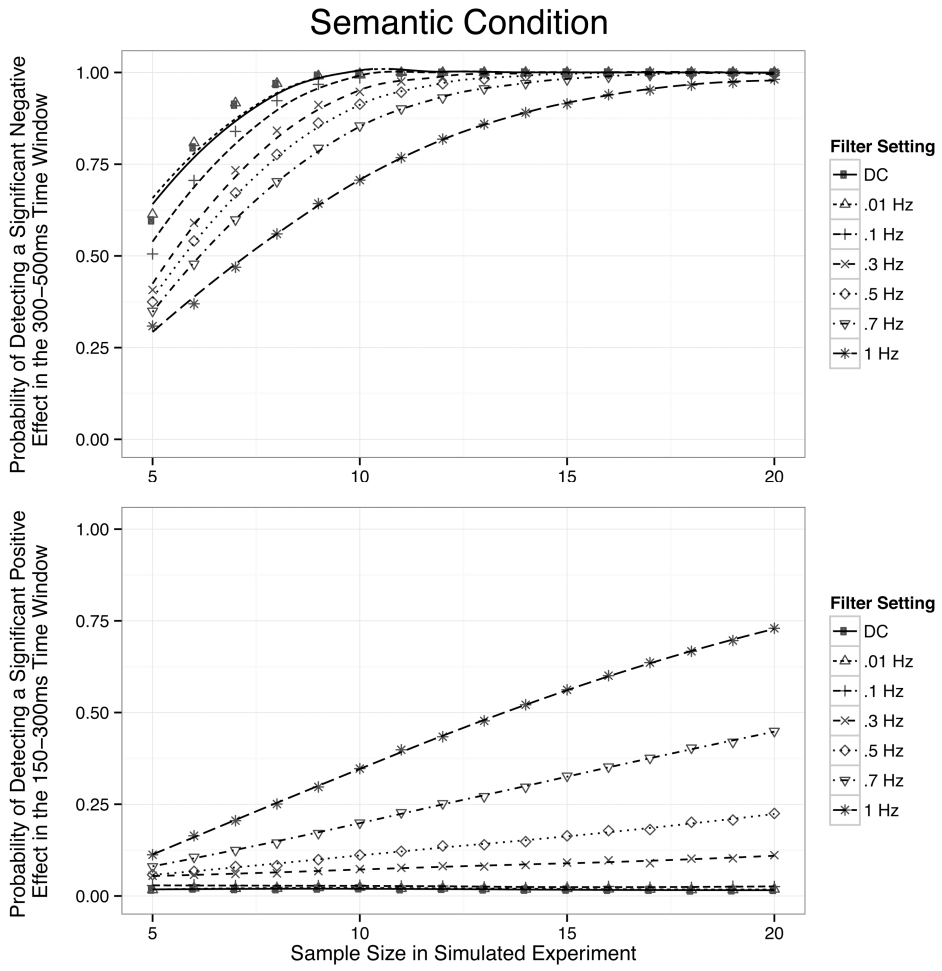
Figure 9.
Probability of detecting significant effects in the semantic condition as a function of sample size and filter setting, as revealed by Monte Carlo simulations. The upper panel depicts power to detect a significant negativity in the 300-500 ms time window (an N400 effect); the lower panel depicts power to detect a significant negativity in the 150-300 ms time window (an artifactual P2 effect). Smooth lines are fit by LOESS.

Table 1

ERP effect magnitudes and one-sample t-statistics as a function of filter setting in the syntactic condition (SEM in parentheses)

	Time Window					
	150-300 ms		**300-500 ms**		**500-800 ms**	
HPF Setting	**Amplitude Difference (μV)**	*t-statistic*	**Amplitude Difference (μV)**	*t-statistic*	**Amplitude Difference (μV)**	*t-statistic*
DC	-0.372 (.392)	-0.952	-0.582 (.462)	-1.259	1.871 (.662)	2.826**
0.01 Hz	-0.329 (.400)	-0.823	-0.513 (.473)	-1.084	1.977 (.660)	2.996**
0.1 Hz	-0.200 (.364)	-0.550	-0.310 (.406)	-0.764	2.292 (.605)	3.793***
0.3 Hz	-0.494 (.338)	-1.465	-0.798 (.339)	-2.357*	1.539 (.501)	3.072**
0.5 Hz	-0.600 (.301)	-1.981+	-1.115 (.289)	-3.863***	0.814 (.356)	2.285*
0.7 Hz	-0.420 (.272)	-1.543	-1.052 (.245)	-4.291***	0.458 (.268)	1.712
1 Hz	-0.723 (.231)	-0.316	-0.696 (.179)	-3.913***	0.327 (.193)	1.697

Note: HPF = high-pass filter.

+
  $p < .06$,

*
  $p < .05$,

**
  $p < .01$,

***
  $p < .001$.

Table 2

ERP effect magnitudes and one-sample t-statistics as a function of filter setting in the semantic condition (SEM in parentheses)

	Time Window					
	150-300ms		300-500ms		500-800ms	
HPF Setting	Amplitude Difference (µV)	t-statistic	Amplitude Difference (µV)	t-statistic	Amplitude Difference (µV)	t-statistic
DC	-0.106 (.341)	-0.310	-2.334 (.372)	-6.282***	-1.781 (.577)	-3.088**
0.01 Hz	-0.091 (.330)	-0.276	-2.313 (.365)	-6.336***	-1.740 (.559)	-3.113**
0.1 Hz	-0.012 (.358)	-0.034	-2.188 (.372)	-5.881***	-1.547 (.570)	-2.714*
0.3 Hz	0.298 (.352)	0.843	-1.717 (.310)	-5.533***	-0.894 (.373)	-2.393*
0.5 Hz	0.491 (.345)	1.424	-1.433 (.272)	-5.265***	-0.639 (.277)	-2.305*
0.7 Hz	0.692 (.324)	2.133*	-1.078 (.220)	-4.890***	-0.348 (.199)	-1.754
1 Hz	0.809 (.281)	2.875**	-0.659 (.158)	-4.173***	0.006 (.139)	.040

Note: HPF = high-pass filter.

+ $p < .06$,

* $p < .05$,

** $p < .01$,

*** $p < .001$.