Investigating Adaptation to *And*-Coordination in English: An ERP Study

Edith Kaan

Department of Linguistics, University of Florida, P.O. Box 115454, Gainesville, FL 32611, USA; kaan@ufl.edu

Abstract: According to certain approaches to adaptation, readers and listeners quickly adjust their processing of sentences to match properties of recently encountered sentences. The present preregistered study used ERP (event-related brain potentials) to investigate how and when readers change their processing in response to recent exposure to sentences of a particular structure. We presented English speakers (*n* = 36) with three virtual blocks of English sentences with *and*-coordination ambiguities. In the first and third block, the ambiguity was always resolved towards a noun phrase (NP-) coordination; in the second block, the structure was always a clausal (S-) coordination. We manipulated the plausibility of the critical noun after the conjunct. N400 and P600 plausibility effects were probed to see to what extent the reader preferred an NP- coordination or expected the sentence to continue differently. Our results suggest that readers change their processing as a function of recent exposure but that they do not immediately adapt to the target structure. Furthermore, we observed substantial individual variation in the type and change in response over the course of the study. The idea that structural adaptation is immediate and a direct reflection of the properties of the recent context therefore needs to be fine-tuned.

Keywords: adaptation; coordination; EEG; ERP; syntactic ambiguity; N400; P600

1. Introduction

Even within the same language community, speakers and writers differ from each other in their word choice and the type of syntactic constructions they tend to use. In daily life, our understanding is hardly disrupted by such variation. This means that readers and listeners have ways to adapt to this variability. One mechanism that has been proposed is that language users keep track of statistical distributions of cues and patterns in the linguistic and non-linguistic context and adapt their processing to accommodate the distributions in the current context. This adaptation takes place through prediction: based on prior information, predictions are made as to what comes next in the sentence. The difference between what is predicted and the actual input is then used to adjust future predictions so as to minimize prediction error and make communication more efficient (e.g., Kuperberg and Jaeger 2016). This continuous updating of distributional information has been hypothesized to play an essential role in language processing and learning (e.g., Chang et al. 2006, 2012; Jaeger and Snider 2013).

Adaptation to speech sounds has been empirically established (see for an overview, Weatherholtz and Jaeger 2016) as has adaptation to a speaker's lexical choice (e.g., Metzing and Brennan 2003). Evidence for adaptation at the level of syntactic comprehension, however, remains unclear. Adaptation in sentence processing has been investigated by exposing readers to sentences with certain structures and recording changes in processing measures as a function of the amount and type of exposure. Some studies report evidence supporting adaptation to syntactic structures, but results have not been systematic (see for an overview Kaan and Chun 2018). Important issues are therefore what is adapted to

during sentence processing, and when, and on the basis of what information adaptation takes place.

One question is what readers are adapting to if they show changes in sentence processing. Are readers changing their processing expectations for a particular syntactic construction (Fine et al. 2013)? Or are they getting used to the (apparent) errors, anomalies or complexities in the experimental materials (Kaschak and Glenberg 2004)? Is the process of dealing with the unexpectancies or anomalies (revision) becoming more efficient (Yan and Jaeger 2020)? Or are readers just getting used to more general aspects of the task (Prasad and Linzen 2019)?

Another question is how quickly and on what basis readers adapt. Do readers adapt their processing on a trial-by-trial basis and are they continuously updating their predictions based on the properties of the recent input (Delaney-Busch et al. 2019; Ness and Meltzer-Asscher 2021)? Or is adaptation slower and based on more general properties of larger chunks of preceding input (Nieuwland 2021a, 2021b)?

A common approach for testing syntactic adaptation in comprehension is to investigate the processing of so-called syntactic garden-path sentences as a function of recent exposure. Garden-path sentences are sentences that contain a temporary syntactic ambiguity which later gets resolved towards the initially non-preferred reading. For example, in *the solders warned…* the verb *warned* can either be a main clause verb (as in *The soldiers warned about the danger*), or can be the verb in a reduced relative clause (as in *The soldiers warned about the danger conducted the raid*). Most English speakers will prefer the more frequent and syntactically simpler main clause reading. When they then see or hear the second verb (*conducted*) that disambiguates towards a reduced relative reading, they experience processing difficulty. The difference in response times to the disambiguating information versus a control in a non-ambiguous context (e.g., *the soldiers that were warned…*) is referred to as the "garden-path effect". Several self-paced reading studies have reported a reduction in the garden-path effect as more sentences of the target structure are read (e.g., Atkinson 2016; Dempsey et al. 2023; Fine et al. 2010, 2013; Kaan et al. 2019). It is tempting to interpret an increasingly smaller garden-path effect as evidence that readers start expecting the initially less-preferred structure. One main concern with this interpretation is that a reduction in the garden-path effect over the course of the study can also be due to task adaptation effects (Prasad and Linzen 2019). As the participants gets used to the task, their self-paced reading gets faster over the course of the experiment. This decrease in response time may be steepest for materials that initially have longer reading times (since the easy items are closer to the maximum response speed to start with). This results in a reduction in the garden-path effect as the experiment progresses, regardless of whether the reader is adapting to the syntactic structure. To reduce such task effects, some researchers use between-participant designs. In these studies, two groups read the same critical items in the second part of the study. In the first part, one group is first exposed to items of the same syntactic structure as the items in the second part; the other group is first exposed to distractor items with a different structure. In this way, both groups have read the same number of sentences in the study by the time they get to the critical second part, but the groups differ in their prior exposure to the critical structure. Some between-group studies find a larger reduction in the garden-path effect for the group that was exposed to more items of the critical structure (Dempsey et al. 2020, Exp. 2 and 3; Fine et al. 2013), some do not (Dempsey et al. 2020, Exp. 1; 2023, Exp. 1; Harrington Stack et al. 2018). In addition, some studies observe a reversal of garden-path effects. These studies have a third part in which the a priori frequent structure (e.g., main clauses) is presented, The group that has read more items of the a priori infrequent structure (e.g., reduced relatives) in previous parts shows longer reading times when the ambiguity is resolved towards the typically frequent structure compared to the control group (Fine et al. 2013). This preference reversal effect is convincing evidence that adaptation has taken place. However, many studies fail to find such reversal effects (c.f., e.g., Atkinson 2016; Dempsey et al. 2020, 2023; Harrington Stack et al. 2018). Power analyses (Prasad and

Linzen 2021) suggest that the lack of consistency among studies may be due to a lack of power, and that more than 1200 participants are needed to detect adaptation in most between-group studies using self-paced reading.

An alternative method to track changes in sentence processing is ERP (event-related brain potentials). Typically, ERP reading experiments use machine-paced, word-by-word presentation of sentences to ease time-locking of the brain signal to the stimuli and to reduce eye movement and other artifacts. This method may therefore diminish task adaptation effects that can be attributed to changes in participants' response times. A handful of ERP studies provide evidence supporting syntactic adaptation. Syntactic violations elicit a larger P600 component relative to their grammatical counterparts. This is a posteriorly distributed positivity, peaking around 500–700 msec after onset of an (apparent) ungrammaticality (e.g., Osterhout and Holcomb 1992). This component has been associated with revision processes (e.g., Friederici 1995; Kuperberg et al. 2020). Several studies manipulated the proportion of sentences with syntactic violations (e.g., 80% or 20% ungrammatical) and investigated how this affected the P600 and other responses (Coulson et al. 1998; Gunter et al. 1997; Hahne and Friederici 1999; Yano et al. 2020). In these studies, the P600 effect was found to be substantially reduced or even numerically reversed (Hahne and Friederici 1999) for the blocks with a high percentage of ungrammatical sentences. Yano et al. (2020) also investigated the modulation of the P600 component over the course of the experimental blocks: by-trial ERPs showed a decrease in the P600 over trials when a high proportion of the sentences contained syntactic violations, but an increase over trials when only a few sentences were ungrammatical. These findings suggest that comprehenders change the way they process syntactic errors based on the frequency with which they encounter such errors in the prior context. However, it is unclear what the comprehender adapts to: a smaller P600 may reflect revision becoming more efficient, syntactic errors becoming more expected, or readers processing the sentences less deeply as they encounter more sentences with syntactic errors.

ERP studies investigating adaptation to semantic anomalies have yielded different results. These studies typically targeted the N400 component. This is a negative-going component, centrally distributed over the scalp, peaking between 300–500 msec after word onset. The N400 is larger for words that are semantically anomalous or less expected given the preceding context (Kutas and Hillyard 1980, 1984). The N400 effect is the difference in amplitude for such words vs. semantically expected and plausible control words. Studies using isolated sentences with semantically unexpected or implausible words did not find a modulation of the N400 effect as a function of the percentage of semantically implausible items (Zhang et al. 2019) or the number of sentences in the prior context with semantic violations (Yano et al. 2020). This suggests that readers do not easily adapt to the occurrence of semantic anomalies in sentences (unless such anomalies are associated with a particular talker or supported by the discourse context; see Boudewyn et al. 2019; Brothers et al. 2019).

On the other hand, ERP studies on related and unrelated word pairs (Lau et al. 2012; Ness and Meltzer-Asscher 2021) report that readers change their semantic processing depending on the prior context. The N400 amplitude is smaller when a word is preceded by a related vs. unrelated word (e.g., Bentin et al. 1985). Lau et al. (2012) found that the N400 effect for related versus unrelated word pairs was larger when the experimental block contained a larger proportion of related word pairs as filler items, suggesting that readers can change their expectation for a related or unrelated word. Delaney-Busch et al. (2019) conducted a by-trial analysis on the Lau et al. data and showed that the trial-by-trial N400 amplitude was a function of the number of preceding trials with related versus unrelated word pairs, among other factors. However, a reanalysis of the data (Nieuwland 2021a) showed no difference in adaptation over the course of the experimental blocks even though the blocks differed in the proportion of related word pairs. This is problematic for the view that readers adapt their processing based on a trial-by-trial tallying of distributional information (see also Ness and Meltzer-Asscher 2021; Nieuwland 2021b).

In sum, prior ERP studies suggest that adaptation to syntactic anomalies is more robust than adaptation to semantic anomalies. Readers change their expectation or processing of syntactic anomalies in response to the frequency of the syntactic violations in the study, but not so much their expectation or processing of semantic anomalies in isolated sentences. Adaptation effects have been found for word pairs, but do not support the view that adaptation closely follows the properties of the recent input.

In the current study, we took a different approach to investigating adaptation to sentence structures using ERPs. Instead of using syntactic violations, we looked at changes in the processing of grammatical sentences that were temporarily ambiguous between two syntactic structures. We probed changes in readers' processing expectations before the point of syntactic disambiguation as a function of recent exposure to one structure or the other. In addition, we looked at changes in processing (revision) at the disambiguating word as a function of exposure.

The Present Study

The overall goal of the present ERP study was to see to how readers change their online processing expectations as a function of recent exposure to particular sentence types. To this aim we presented participants with three virtual blocks of sentences (see Table 1) with *and*-coordinations. These constructions are temporarily ambiguous since *and* can connect two noun phrases (as in a and b in Table 1), two clauses (as in c and d), or other syntactic phrases. Prior research has shown that readers prefer the interpretation in which *and* conjoins two noun phrases (as in a, b) when no context is provided (e.g., Engelhardt and Ferreira 2010; Frazier 1987; Hoeks et al. 2006; Kaan and Swaab 2003; Staub and Clifton 2006). The experiment consisted of three virtual blocks. In the first block, we presented readers with noun phrase (NP-)coordinations, which is the preferred resolution of *and* coordinations. In the second block, we changed the structure to clausal (S-)coordinations (as in c, d). In the last block, we changed the structure back to the NP-coordination (e, f). In order to track participants' processing expectations, we included a plausibility manipulation: the noun after *and* was either plausible as the object of the first verb (b, d, f), or implausible (a, c, e).

Table 1. Examples of the experimental conditions.

Block	Condition [1]	Sentence [2]
1	a. NP-coordination, implausible	The child ate the popcorn and the <u>*movie*</u> about the *superhero*.
1	b. NP-coordination, plausible	The child enjoyed the popcorn and the <u>*movie*</u> about the *superhero*.
2	c. S-coordination, implausible	The child ate the popcorn and the <u>*movie*</u> about the superhero <u>made</u> him happy.
2	d. S-coordination, plausible	The child enjoyed the popcorn and the <u>*movie*</u> about the superhero <u>made</u> him happy.
3	e. NP-coordination, implausible	The child ate the popcorn and the <u>*movie*</u> about the *superhero*.
3	f. NP-coordination, plausible	The child enjoyed the popcorn and the <u>*movie*</u> about the *superhero*.

[1] "Plausible" and "implausible" refers to the relation between the critical noun and the preceding verb. [2] Critical noun is underscored and in italics; last words analyzed are in italics; verbs analyzed are underscored. Markings are for illustration purposes only and were not seen by the participants.

Our predictions were as follows. First, we expected changes in the ERPs between the three virtual blocks at the noun after *and*, henceforth the "critical noun". If readers prefer an NP-coordination, they should show effects of semantic implausibility at this noun. We therefore expected the ERPs to the critical noun in the first block to show a difference

between the conditions in which the noun was implausible vs. plausible as the object of the preceding verb. In the second block, *and* connects two clauses, and the noun after *and* is no longer the object of the preceding verb. Instead, it is the subject of the following verb. If readers adjust their preferences and start to expect an S-coordination, the critical noun phrase should not be interpreted as the object of the preceding verb; implausibility effects at the noun after *and* in (c) versus (d) should be smaller than in (a) versus (b). In the final block of the study, we presented participants again with NP-coordinations. If readers change their strategy again, we expect the implausibility effects at the noun after *and* to be larger than in block 2.

In particular, we expected a modulation of ERP components commonly observed in response to semantic anomalies: the N400 component and late positivities. As mentioned above, the N400 is a centrally distributed negativity, peaking around 300–500 ms after onset of a content word. It is larger for words that induce a semantic violation or are unexpected compared to words that are predicted and plausible. One interpretation is that the N400 reflects the activation of information that has not already been pre-activated by the context (Kuperberg et al. 2020). For instance, an expected word can be almost completely pre-activated and elicits a small or no N400. A word that is unexpected has not many features pre-activated by the context, and hence elicits a large N400. Semantically unexpected words can also elicit later positive-going components (around 500–900 ms). A later frontal positivity (Kuperberg et al. 2020; Van Petten and Luka 2012) has been observed when an incoming word is plausible given the context but does not match the word that was highly predicted. The reader then needs to update the context to fit in the new information. A posterior positive-going component (P600) has been found for input that is, or seems to be, semantically or syntactically anomalous, and that cannot be integrated into the (event-semantic, syntactic, or higher-level) representations maintained thus far. This component may reflect revision and repair (Kuperberg et al. 2020). We therefore probed the N400 as well as later positivities at the critical noun.

Our second prediction pertained to the last word of the sentence in (a-b) and (e-f). Plausible sentences typically elicit a central-posterior end-of-sentence positivity (Kutas and Hillyard 1983; Van Petten and Kutas 1991); sentences containing a mid-sentence semantic anomaly can elicit a sentence-final N400 (De Vincenzi et al. 2003; Osterhout and Nicol 1999; see for an overview Stowe et al. 2018). We therefore expected a sentence-final negativity for (a) versus (b) in block 1. If readers adapt to *and* connecting two clauses in block 2, they will continue to predict a second verb at the start of block 3. The end of the sentence in block 3 will then initially be highly unexpected. Prior studies reported an N400-like negativity at the end of sentences that can be interpreted as syntactically incomplete (Osterhout and Holcomb 1992, 1993). We therefore expected that the ERPs to the sentence's final word in block 3 would be more negative overall than in block 1. Furthermore, if readers change their preference to an NP-coordination over the course of block 3, a plausibility effect may be seen, as in block 1, but this is expected to be smaller for the entire block given the initial preference for an S-coordination at the start of block 3.

Finally, we analyzed the ERPs at the second verb in (c) and (d). Prior research (Kaan and Swaab 2003) has shown a P600 component at the disambiguating verb in constructions such as (f) versus a non-ambiguous condition. Based on studies using slightly different ambiguities (Román et al. 2021), we expected the P600 at the verb to be initially larger for conditions in which the critical noun phrase is plausible as a direct object of the preceding verb. This is because revision of the structure is easier if there is a semantic bias away from the initial reading (*and* connecting two noun phrases in this case). If readers start to expect in block 2 that *and* connects two clauses, the P600 should become smaller overall over the course of block 2, and less affected by the plausibility of the noun after *and*.

The above predictions were based on the assumption that adaptation takes place within a few trials (Fine et al. 2013) and hence, ERPs for each block on average will be different. To further assess more fine-grained changes in processing, we also conducted

exploratory analyses in which we looked at by-trial changes in the plausibility effect at the noun and last word within a block.

As a preview of our results, our findings suggest that readers adapt their processing to properties of the preceding sentence structures but that some aspects of processing are slower to change than others. Furthermore, some aspects of processing do not reflect the local frequencies of the sentence types. We also observed a large variation between individuals in the type of response and changes therein over the course of the study.

2. Materials and Methods

The methods and analyses were preregistered on OSF (https://osf.io/szgfd/registrations.)

2.1. Participants

Data were collected from 42 participants recruited from the University of Florida community. Participants received a USD 10/hour monetary compensation or additional course credit. Data from 6 participants were omitted from the analysis because of technical problems (2 participants); excessive blinking or other artifacts leading to fewer than 18 artifact-free trials per condition (1); participants falling asleep during the study (2); or not completing the study (1). The remaining 36 participants (which was the preregistered total of datasets; 24 women, 12 men, age 18–25 years, mean age 20.0) were all right handed and reported growing up in a monolingual English speaking household. Of these, 23 participants indicated having learned another language (mean age of acquisition: 13.1; mean self-rated proficiency: 2.8 out of 7; mean percentage of current exposure to the second language: 4.2%). Participants had no history of reading or learning disorders as determined by self-report.

2.2. Materials

The experiment consisted of three virtual blocks of two conditions each. The conditions are illustrated in Table 1. Sentences in all conditions contained the conjunct *and* which was preceded and followed by a noun phrase. The noun phrase following *and* was either a plausible object of the preceding verb (b, d, f), or was implausible as the object of the first verb (a, c, e). In the first block of the study, *and* was used as a noun-phrase coordination. In the second block of the study, *and* connected two clauses. Sentences in block 2 started out in the same way as in block 1, but continued with an inflected verb and other material. In the last block, block 3, *and* again connected two noun phrases. The noun after *and* was always followed by a phrase of at least 2 words. This was to avoid disambiguation by punctuation at or right after the critical word. For conditions (c) and (d) in block 2 at least 2 words followed the disambiguating verb (e.g., *made* in Table 1).

Materials Norming

Plausibility rating: To check that the critical noun was implausible as the object of the first verb in conditions (a), (c), and (e), and plausible in conditions (b), (d), and (f), we conducted a plausibility rating using Mechanical Turk. Sentences in the NP-coordination plausible and implausible conditions were presented up to and including the noun after *and*, and ended in a period (e.g., *The child ate the popcorn and the movie*). We collected data from 300 participants (age 19–74, mean age 37.4), restricting the IP address to the US. Each participant saw 40 experimental items (20 designed to be plausible, 20 designed to be implausible), plus 44 distractor items of varying plausibility, and rated these on a Likert scale from 1 ("very implausible/very unlikely") to 7 ("very plausible/very likely.") We excluded the responses of 51 users who either indicated having learned English after the age of 16, gave the same response for most trials, or gave an average rating of higher than 5 to the 4 filler items that were designed to be very implausible. For the remaining data, we averaged the scores assigned to each sentence (19 to 24 responses per item). On the basis of

these plausibility scores, we selected 216 item sets to be used in the main experiment. On average the versions intended to be implausible received a plausibility score of 3.14, range 1.88–4.38), the versions intended to be plausible received a score of 5.77 (range 4.90–6.50; $T(215) = 65.35$, $p < 0.0001$).

Association norming: We conducted another study using Mechanical Turk to obtain a measure of the association between the nouns on either side of *and*. We collected data from 60 participants (age 23–66, average 39.9), restricting the IP address to the US. Participants used a 7-point Likert scale (1: not associated at all, to 7: very strongly associated) to rate the 216 noun pairs used in the study (e.g., *milk AND sandwich*), plus 6 distractor items designed to be non-associated. The nouns were presented as bare nouns and appeared in the same position around *and* as in the experimental items. As in the plausibility norming study, we dropped data from those participants who indicated having learned English after the age of 16, from those who consistently answered using one particular score, and from those who rated the 6 non-associated distractor items on average higher than 5. For each item, the average rating of the remaining 47 participants was used as an index of association of the two nouns (mean association score: 4.32; *SD* 1.22; range 1.49–6.57). None of the items were dropped on the basis of the association scores.

For the main study, the 216 sextuplets were divided into six matched groups of 36 items using Match (van Casteren and Davis 2007). The six groups did not statistically differ on the following measures: plausibility ratings for the plausible and implausible versions; noun association ratings; the length and frequency of the noun after *and*; the length and frequency of the disambiguating verb in the S-coordination conditions; and the length and frequency of the sentence-final word in the NP-coordination conditions. Frequency was the log transformed lemma frequency per million extracted from the British National Corpus (BNC Consortium 2007) using the NIM search engine (Guasch et al. 2013). Six participant lists were then created by Latin squaring the item groups such that all items from a group appeared in one condition on a particular list, and in another condition on another list. In this way each participant list contained 36 items for each of the 6 conditions, no item was repeated within a list, and each item was presented in all conditions across the experiment as a whole. An additional 84 plausible filler sentences of various structures were included (same for each list) to add variation. These filler items were on average of the same length as the experimental items (e.g., *The train departed from the station that was on the north side of town*). Twenty-eight percent of the items were followed by a comprehension question (60 of the 216 experimental items; 24 of the 84 distractor items). The questions mainly probed the first clause, and not the resolution of the ambiguity of *and* (example: *The child enjoyed the popcorn and the movie about the superhero*. Question: *Did the child eat a salty snack?* (*Yes*)). The complete set of materials and norms are available at https://osf.io/szgfd/.

2.3. Procedure

The participants were seated in a chair, about 1 m from a computer monitor in an electrically shielded sound-attenuating booth. Participants first filled out in-house questionnaires about health, handedness, and language background, and were fitted with an electrode cap. In the reading experiment, stimuli were displayed and behavioral responses were collected using Eprime 2.0 pro (Psychological Software Tools, Pittsburgh, PA). Sentences were presented one word at a time, white font on black background. Each word was presented for 300 msec with a 200 msec blank screen separating the words. Before each sentence, a fixation cross appeared on the screen for 500 msec. After the final word of each sentence, the screen remained blank for 1000 msec. Then, either the message "Press for Next" or a comprehension question was presented. Participants used a game pad to answer by pressing the right button for "yes", and the left button for "no". Participants were instructed to silently and attentively read the sentences from the screen, while trying to refrain from blinking while the sentences were presented.

The 300 sentences per participant list were divided into 10 runs each consisting of 30 sentences. The transition between the first (conditions a and b) and second blocks (conditions c and d) of the study, and between the second (c and d) and final blocks (e and f) of the study occurred midway through runs 4 and 7, respectively. This was to make the participant less aware of the virtual segmentation and to maximize the effects of the change in structure. Participants were not told that there were differences between the runs. The order of the runs within each block (runs 1–3; 5–6; 8–10) was randomized between participants. In addition, the item order was different for each of the six participant lists. Participants were presented with 8 practice sentences (3 followed by comprehension questions) before the experimental set. Right after electrode application and before the sentence reading experiment, we collected resting-state EEG data (3.5 min of eyes open, 3.5 min of eyes closed). We will not report these resting-state data in this paper. The entire testing session lasted about 2.5 h per participant.

2.4. EEG Recording

We recorded from 64 Ag/AgCl electrodes mounted in an elastic cap (ANT-Neuro Waveguard TM). Additional electrodes were placed above and below the right eye, and on the outer canthi to monitor EOG. Another set of electrodes were placed on the left and right mastoid processes. EEG was collected using an ANT Refa 78 amplifier (ANT-Neuro, Hengelo, The Netherlands) at a rate of 512 Hz, referenced to AFz.

2.5. Data Analysis

Data preprocessing and analysis was conducted using EEGlab (Delorme and Makeig 2004) and ERPlab (Lopez-Calderon and Luck 2014) running in MATLAB. The signal was re-referenced off-line to the mean of the left and right mastoid, and first filtered between 0.1 and 55 Hz.[1] Eye movements were corrected using ICA (Jung et al. 2000).

Next, data were filtered between 0.1 and 30 Hz. Epochs were defined from −200 to 1000 ms relative to the onset of the critical noun (noun after *and*), disambiguating verb (in conditions c and d) and last word (conditions a–b, e–f). Trials with excessive deflections (containing an absolute voltage over 75 μV; exceeding a peak-to-peak threshold of 60 μV at AF1/z/2 in a 200 ms time window (100 ms steps), or a step-like artifact exceeding a 40 μV threshold in a 200 ms time window (50 ms steps)) were automatically flagged and rejected from the analysis. Bad channels were identified by visual inspection and interpolated. The resulting mean number of trials per condition ranged from 30.6 to 32.9 for the critical nouns in the six conditions; 32.2 and 32.8 for the verbs in block 2; and ranged from 32.3 to 32.7 for the sentence-final words in blocks 1 and 3. All participants had at least 20 resulting trials per condition.

For each participant, we obtained the mean amplitude in critical time windows (defined below) for each electrode and trial. The N400 was quantified as the mean amplitude between 300 and 500 msec after target word onset across the following electrodes: Cz, C1, C2, C3, C4, CPz, CP1, CP2, CP3, CP4; and the P600 as the mean amplitude between 500 and 900 msec after onset averaged across the following electrodes CPz, CP1, CP2, CP3, CP4, Pz, P1, P2, P3, P4. We quantified the frontal positivity as the mean amplitude between 500 and 900 msec across Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4. However, no frontal effects were found. We will therefore not report results from these analyses below. The 200 ms pre-stimulus interval was used as a baseline. For each question of interest, we estimated a linear mixed-effects model using the lme4 packet version 1.1-21 (Bates et al. 2015b) in R version 4.1.2 (R Core Team 2021). We preregistered analyses for the epochs pertaining to (1) the noun phrase after *and*; (2) the sentence-final word in the first and third block; and (3), the verb in the second clause in the second block. For (1), plausibility (treatment coded, with plausible as the reference level), block, and their interaction were entered as fixed effects. Block was treatment coded with block 2 as the reference level. Hence, block 1 was compared to block 2, and block 3 was compared to block 2. For (2), the model

contained plausibility, block (block 1 and block 3) and their interactions. Plausibility and block were treatment coded, with plausible and block 1 as reference level. For (3), plausibility (with plausible as reference level) and trial position (position of the trial in the block (centered) relative to other experimental trials) were included as fixed effects. In all models, participant and item were included as random intercepts. We started with a model in which all fixed effects and their interactions (where applicable) were included as random slopes. Random effects were eliminated starting with factors that had the smallest variance until the model converged without warnings (Bates et al. 2015a). In all cases, our final models contained only by-participant and by-item intercepts as random effects. *p*-values were obtained based on Satterthwaite's method using LmerTest version 3.1-0 (Kuznetsova et al. 2017). For the complete models and results see the Supplementary Materials, and https://osf.io/szgfd/.

3. Results

3.1. Question Answering

On average, our participants responded to the comprehension questions with a 92.8% accuracy (*SD* = 4.2%, range 81–99%). This suggests they were paying attention to the sentences. Since only 28% of the sentences were followed by a question, we included all trials in the EEG analysis regardless of the participants' accuracy on the trials. Question answering accuracy did not significantly differ in the first three runs compared with the last three runs of the study, suggesting participants were still paying attention towards the end of the study (mean accuracy (*SD*), first three runs: 94.3% (4.2), last three runs: 92.9% (6.8); *Estimate* = −0.23, *SE* = 0.19, *z* = −1.20, *p* = 0.23).

3.2. ERPs, Preregistered Analyses

3.2.1. Noun after *and*

Figure 1 displays ERPs for the N400 and P600 electrode regions for each of the three blocks. We had predicted plausibility effects (that is, the differences between plausible and implausible conditions in N400 or late positive components) to be smaller in block 2 (S-coordinations) than in blocks 1 and 3 (NP-coordinations). However, the N400 plausibility effect was numerically *larger* in block 2 compared with blocks 1 and 3 (mean plausibility effect: block 1, −0.30 μV (*SD* = 1.69); block 2, −0.69 μV (*SD* = 1.97); block 3, −0.08 μV (*SD* = 2.31)).

The outcomes of the linear mixed-effects model for the N400 effect are given in Table 2. Using block 2 and the plausible condition as the reference level, the effect of plausibility on the N400 was significant, with the implausible condition showing a larger negativity (plausibility: *estimate* = −0.68; *95% CI* [−1.30, −0.07]; *SE* = 0.31; *T* = −2.19; *p* = 0.03). Plausibility by block interactions were not significant. The plausibility effect, however, disappeared when block 1 was the reference level (*estimate* = −0.30, *95% CI* [−0.92, 0.31], *SE* = 0.31, *T* = −0.96, *p* = 0.34), and when block 3 was the reference level (*estimate* = -0.08, 95% CI [−0.70, 0.53], *SE* = 0.32; *T* = −0.27, *p* = 0.79), see Supplementary Tables S1 and S2. The N400 plausibility effect was therefore robust only in block 2.[2]

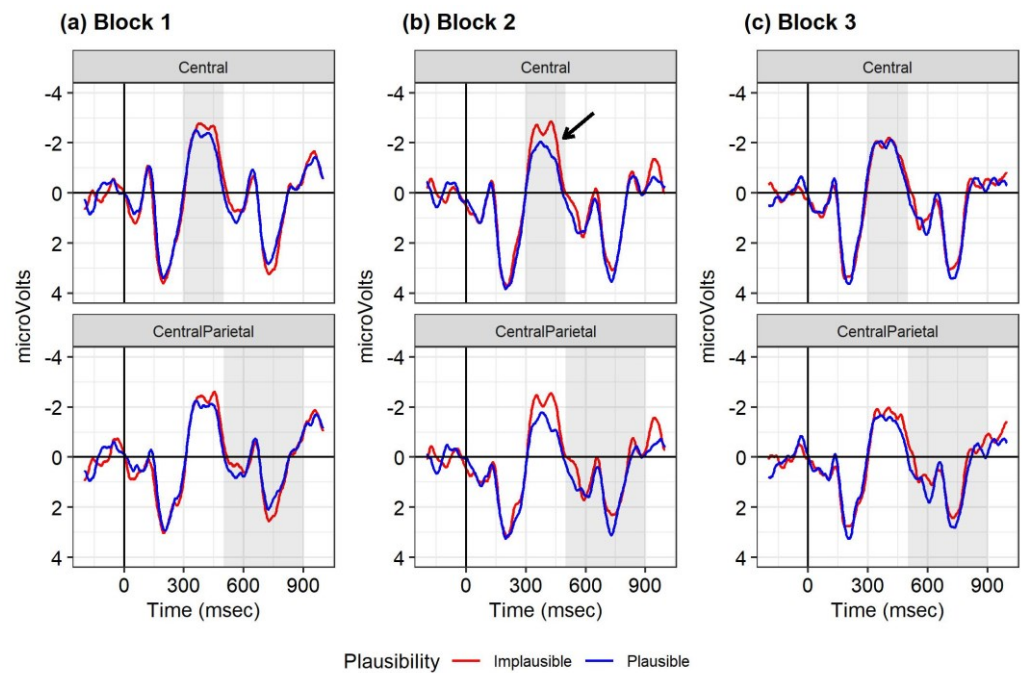The P600 analysis did not yield any significant effects; see Table 3.

Figure 1. ERPs at the critical noun for the central region (top row) and centro-parietal region (bottom row), averaged for the electrodes in each region. Negative polarity goes upwards on the y-axis. Red line: conditions in which the noun is implausible given the preceding verb; blue line: conditions in which the noun is plausible given the preceding verb. Shading indicates the time window analyzed. Arrow indicates the N400 plausibility effect in block 2; (**a**) block 1 (NP-coordinations); (**b**) block 2 (S-coordinations); (**c**) block 3 (NP-coordinations).

Table 2. Results of the linear mixed-effects model of the N400 effect at the critical word (300–500 msec, central electrodes).

Parameter [1]	Estimate	95% CI	SE	T	p
Intercept	**−1.27**	**−1.84, −0.71**	**0.29**	**−4.43**	**<0.001** [2]
Plausibility	**−0.68**	**−1.30, −0.07**	**0.31**	**−2.19**	**0.029**
Block 1 vs. 2	−0.52	−1.13, 0.09	0.31	−1.66	0.097
Block 3 vs. 2	−0.31	−0.92, 0.31	0.31	−0.98	0.325
Plaus: block 1 vs. 2 [3]	0.38	−0.48, 1.25	0.44	0.87	0.386
Plaus: block 3 vs. 2	0.60	−0.27, 1.47	0.44	1.35	0.176

[1] Random variance: by-participant intercept 1.07; by-item intercept 0.88; [2] Effects in bold are effects with *p*-value <0.05; [3] Plaus: plausibility.

Table 3. Results of the linear mixed-effects model of the P600 effect at the critical word (500–900 msec, central-parietal electrodes).

Parameter [1]	Estimate	95% CI	SE	T	p
Intercept	**1.10**	**0.59, 1.61**	**0.26**	**4.22**	**<0.001**
Plausibility	−0.28	−0.88, 0.31	0.30	−0.94	0.348
Block 1 vs. 2	−0.58	−1.17, 0.01	0.30	−1.91	0.056
Block 3 vs. 2	−0.11	−0.70, 0.49	0.30	−0.35	0.728
Plaus: block 1 vs. 2	0.35	−0.49, 1.19	0.43	0.83	0.409
Plaus: block 3 vs. 2	−0.03	−0.88, 0.81	0.43	−0.08	0.937

[1] Random variance: by-participant intercept: 0.72; by-item intercept: 0.43.

3.2.2. Last Word in NP-Coordinations (Blocks 1 and 3)

ERPs to the sentence-final word in blocks 1 and 3 are displayed in Figure 2. Outcomes of the analysis of the N400 and P600 regions are given in Tables 4 and 5, respectively. We had expected that the ERPs at the end of the sentence in block 1 would be more negative/less positive for the implausible condition. In addition, if readers adapt to *and* connecting two clauses in block 2, the end of the clause in block 3 would initially be unexpected to the readers. We therefore expected ERPs at the end of the sentence to be more negative overall in block 3 than block 1, with a smaller plausibility effect compared with block 1. The first two predictions were borne out. A linear mixed-effects model with block 1 and the plausible condition as reference levels showed an effect of plausibility for the P600 region of interest: the ERPs were more positive for the plausible than the implausible condition (*estimate* = −0.74, *95% CI* [−1.37, −0.12], *SE* = 0.32, *T* = −2.33, *p* = 0.02). Furthermore, the sentence-final ERPs in block 3 were overall more negative than in block 1 in the N400 and P600 analyses (effect of block: N400: *estimate* = −0.69, *95% CI* [−1.33, −0.04], *SE* = 0.33, *T* = −2.08, *p* = 0.04; P600: *estimate* = −0.95, *95% CI* [−1.57, −0.33], *SE* = 0.32, *T* = −2.99, *p* = 0.003). The interaction between plausibility and block was not significant.
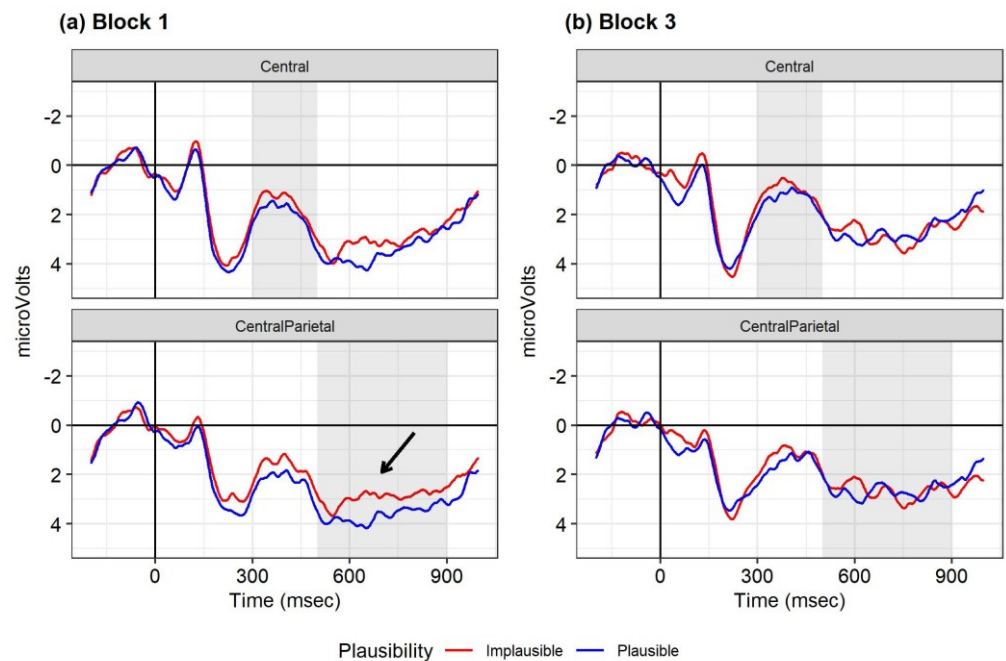

Figure 2. ERPs at the sentence-final word in the NP-coordinations for the central and centro-parietal regions, averaged for each of the electrode regions analyzed. Negative polarity goes upwards on the y-axis. Red line: conditions in which the critical noun is implausible given the preceding verb; blue line: conditions in which the critical noun is plausible given the preceding verb. Shading indicates the time window analyzed. Arrow indicates the sentence-final plausibility effect in block 1. (**a**) Block 1; (**b**) block 3.

Table 4. Results of the linear mixed-effects model of the N400 at the sentence-final word in blocks 1 and 3 (300–500 msec, central electrodes).

Parameter [1]	Estimate	95% CI	SE	T	*p*
Intercept	**2.03**	**1.25, 2.80**	**0.40**	**5.13**	**<0.001**
Plausibility	−0.42	−1.07, 0.23	0.33	−1.27	0.205
Block	**−0.69**	**−1.33, −0.04**	**0.33**	**−2.08**	**0.037**
Plausibility: Block	0.09	−0.83, 1.01	0.47	0.19	0.847

[1] Random variance: by-participant intercept: 3.33; by-item intercept: 2.05.
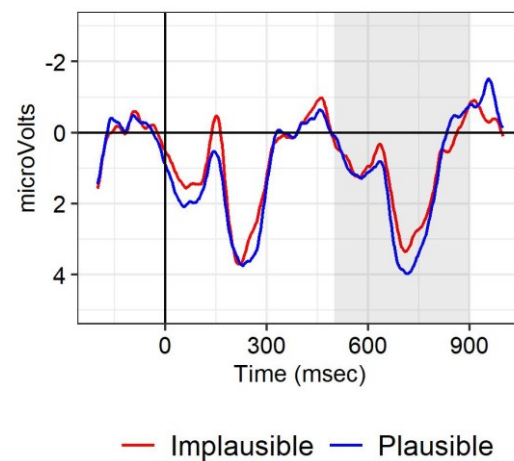
Table 5. Results of the linear mixed-effects model of the P600 at the sentence-final word in blocks 1 and 3 (500–900 msec, central-parietal electrodes).

Parameter [1]	Estimate	95% CI	SE	T	p
Intercept	**3.63**	**2.86, 4.41**	**0.39**	**9.21**	**<0.001**
Plausibility	**−0.74**	**−1.37, −0.12**	**0.32**	**−2.33**	**0.020**
Block	**−0.95**	**−1.57, −0.33**	**0.32**	**−2.99**	**0.003**
Plausibility: Block	0.71	−0.17, 1.59	0.45	1.58	0.115

[1] Random variance: by-participant intercept: 3.40; by-item intercept: 2.35.

3.2.3. Verb in S-Coordinations (Block 2)

Figure 3 depicts the ERPs at the disambiguating verb in block 2. We had expected that the P600 would be larger for conditions in which the noun after *and* was plausible given the preceding verb. Numerically, this pattern was borne out. However, statistical analyses yielded no effect of plausibility or a change of the plausibility effect over the course of block 2 (see Tables 6 and 7).[3]

Figure 3. ERPs at the disambiguating verb in block 2 (S-coordinations) for the centro-parietal region, averaged for the electrode in the region analyzed. Negative polarity goes upwards on the y-axis. Red line: condition in which the critical noun is implausible given the preceding verb; blue line: condition in which the critical noun is plausible given the preceding verb. Shading indicates the time window analyzed according to the preregistered analysis.

Table 6. Results of the linear mixed-effects model of the N400 at the second verb in block 2 (300–500 msec, central electrodes).

Parameter [1]	Estimate	95% CI	SE	T	p
Intercept	0.39	−0.18, 0.96	0.29	1.34	0.181
Plausibility	−0.14	−0.78, 0.51	0.33	−0.41	0.678
Trial position	−0.01	−0.04, 0.01	0.01	−1.19	0.233
Plausibility: Trial position	0.01	−0.03, 0.04	0.02	0.44	0.658

[1] Random variance: by-participant intercept: 0.98; by-item intercept: 1.0.

Table 7. Results of the linear mixed-effects model of the P600 at the second verb in block 2 (500–900 msec, central-parietal electrodes).

Parameter [1]	Estimate	95% CI	SE	T	p
Intercept	**1.48**	**0.83, 2.12**	**0.33**	**4.51**	**<0.001**
Plausibility	−0.23	−0.84, 0.37	0.31	−0.76	0.450
Trial position	−0.01	−0.03, 0.01	0.01	−0.73	0.468
Plausibility: Trial position	0.01	−0.02, 0.04	0.02	0.58	0.564

[1] Random variance: by-participant intercept: 1.68; by-item intercept: 2.83.

3.3. Interim Discussion

We observed an N400 plausibility effect at the critical noun. However, in contrast to our preregistered prediction, this effect was numerically largest in block 2 (S-coordinations) compared to the other blocks (NP-coordinations). In block 1, the sentence-final words showed an increased positivity for the plausible vs. implausible conditions. Plausible sentences typically elicit a sentence-final positivity (Stowe et al. 2018; Van Petten and Kutas 1991). This difference between the implausible and plausible conditions suggests that the readers do notice the implausibility in block 1. ERPs to the sentence-final word were overall more negative in block 3 than in block 1. This suggests that the sentence in both the plausible and implausible condition in block 3 was considered incomplete, or anomalous to some extent compared to block 1.

Taken together, these data suggest that our readers changed their processing in response to recent exposure, but did not immediately change their expectation of the structures to mirror the distribution of these structures in the recent input. If they had, block 2 would have shown no, or a smaller N400 plausibility effect at the critical word, since all sentences in block 2 are globally plausible when processed as an S-coordination.

To further explore changes in processing over the course of the study, we conducted analyses within each block, investigating the effect of the position of the trial relative to other critical trials. We also explored differences in responses to the critical words between individuals.

3.4. Exploratory Analyses

3.4.1. Within-Block Analysis of Trial Position

For each block, we conducted a linear mixed-effects model on the N400 and P600 with plausibility (plausible coded as reference level) and trial position relative to other experimental trials (centered) as fixed effects. The random effects were by-participant and by-item random intercepts. Changes over trial position within a block were attested only at the sentence-final word in block 1 (NP-coordinations); see Tables 8 and 9 (for the results of the other analyses see Supplementary Tables S6–S13). ERPs at the sentence-final position became less positive for the plausible condition over the course of block 1 in the N400 and P600 regions analyzed (N400: estimate = −0.02, 95% CI [−0.05, −0.00], SE = 0.01, T = −2.03, p = 0.04; P600: estimate = −0.02, 95% CI [−0.05, −0.00], SE = 0.01, T = −1.98, p = 0.05). This led to an increasingly smaller plausibility effect as more experimental sentences had been seen in block 1 (plausibility x trial position: N400: estimate = 0.03, 95% CI [0.00, 0.07], SE = 0.02, T = −2.07, p = 0.04, P600: estimate = 0.03, 95% CI [−0.00, 0.06], SE = 0.02, T = 1.69, p = 0.09).

Table 8. Results of the linear mixed-effects model of the N400 at the sentence-final word in block 1 (300–500 msec, central electrodes).

Parameter [1]	Estimate	95% CI	SE	T	*p*
Intercept	2.04	1.20, 2.88	0.43	4.78	<0.001
Plausibility	−0.43	−1.07, 0.20	0.32	−1.34	0.180
Trial position	−0.02	−0.05, −0.00	0.01	−2.03	0.042
Plausibility: Trial position	0.03	0.00, 0.07	0.02	2.07	0.039

[1] Random variance: by-participant intercept: 4.43; by-item intercept: 1.61.

Table 9. Results of the linear mixed-effects model of the P600 at the sentence-final word in block 1 (500–900 msec, central-parietal electrodes).

Parameter [1]	Estimate	95% CI	SE	T	*p*
Intercept	3.64	2.89, 4.39	0.38	9.55	<0.001
Plausibility	−0.73	−1.34, −0.12	0.31	−2.34	0.020
Trial position	−0.02	−0.05, −0.00	0.01	−1.98	0.048
Plausibility: Trial position	0.03	−0.00, 0.06	0.02	1.69	0.090

[1] Random variance: by-participant intercept: 3.04; by-item intercept: 2.75.

3.4.2. Exploring Individual Differences

We did not find strong plausibility effects at the critical noun in blocks 1 and 3. This could be due either to most participants not showing a difference between the plausible and implausible conditions, or to some participants showing a positivity and some a negativity in response to the anomaly, with the net effect that these effects cancel each other out at the group level. (Grey 2022; Pélissier 2020; Tanner and Van Hell 2014). A study on adaptation to coordination ambiguities in second-language learners has also reported individual differences (Kaan et al. forthcoming). To further explore this issue, we computed the response magnitude and response dominance index for the plausibility effect at the critical noun in each block for each participant (Pélissier 2020). The response magnitude is computed by subtracting the N400 amplitude in the implausible condition from that in the plausible condition, and subtracting the P600 amplitude in the plausible from the implausible conditions. Both these values are then squared and summed. The response magnitude (RM) is the square root of this sum. The higher the RM value, the larger the plausibility effects are (N400, P600, or both). The response dominance index (RDI) is obtained by subtracting the magnitude of the N400 plausibility effect (N400 to the plausible condition minus N400 to the implausible condition) from the magnitude of the P600 plausibility effect (P600 to the implausible minus P600 to the plausible condition), and dividing this difference by the square root of 2. Positive RDI values indicate a predominantly positive response; negative RDI values indicate a predominantly negative response. The mean RDIs and RMs for the critical noun over the three blocks are given in Table 10. Note that the RMs are not close to zero and that the variation in RDI is substantial. This suggests that our participants did respond to the implausibility at the noun in blocks 1 and 3, but that the individual differences in the type of response cancelled out the group-level effects. Individuals also differed in how the RDI changed between the blocks. Focusing on the 21 participants who predominantly showed an N400 in block 2 (negative RDI), 4 consistently showed an N400 over all three blocks; 5 showed an N400 in blocks 1 and 2, but changed to a positivity in block 3; 12 changed from a positivity in block 1 to a negativity in block 2. Of these 12, 5 continued to show a negativity in block 3, whereas 7 changed back to a positivity. This suggests that participants also differed in how they adjusted their processing over the course of the study.

Table 10. Mean response magnitude (RM) and response dominance index (RDI) values for the ERPs at the critical noun in the three blocks (standard deviation in parentheses).

Block	Mean RM (SD)	Mean RDI (SD)
Block 1 (NP-coordination)	2.18 (1.20)	−0.16 (2.32)
Block 2 (S-coordination)	2.51 (1.59)	−0.69 (2.66)
Block 3 (NP-coordination)	2.68 (1.74)	−0.28 (3.04)

4. Discussion

The aim of the current study was to investigate changes in processing in response to the type of sentences in the experimental context. We presented readers with temporarily ambiguous *and*-coordinations that were resolved towards a (typically preferred) NP-coordination in the first block of the study, towards an S-coordination in the second block, and again towards an NP-coordination in the third block. We investigated adaptation by looking at ERPs to nouns that were semantically anomalous under the NP-coordination interpretation (hence, expected to elicit an N400 or other plausibility effect), and semantically plausible under the S-coordination interpretation (hence expected to elicit a smaller plausibility effect if the reader is expecting this syntactic structure). In addition, we investigated effects at the sentence-final words and at the disambiguating verb.

Our main findings were the following. First, an N400 plausibility effect was observed at the critical noun. This effect was numerically largest in block 2 (S-coordinations), where we had predicted the plausibility effect to be smallest. The difference in the N400 plausibility effect was not statistically different between the blocks, however. Second, the sentence-final word in block 1 (NP-coordinations) showed a positivity for the plausible versus implausible condition. This plausibility effect became smaller as more NP-coordinations had been seen. Third, the ERPs to the sentence-final word in the NP-coordinations were overall more negative in block 3 than block 1. Finally, we observed a large individual variation in the type of plausibility effects at the critical noun and changes in these effects between the blocks.

Below we will elaborate on and interpret our findings, and discuss our findings in the light of our research question.

4.1. Interpretation of the Findings

Our results for the critical noun were rather unexpected, since our prediction of finding the largest plausibility effects in blocks 1 and 3 was not borne out. We did find a plausibility effect at the sentence-final word in block 1 (NP-coordinations), which decreased in size over the course of the block. An explanation of these results is that our readers did not always strongly prefer *and* to connect two noun phrases, but that this preference was immediately modulated by plausibility. This parallels observations in an eye-tracking study on *and*-coordinations conducted by Hoeks et al. (2006). This study tested only S-coordinations, but, as in the present study, manipulated the semantic fit between the first verb and the noun following *and*. In this paradigm, no plausibility effects were observed in the eyetracking data at the critical noun. Our results are also reminiscent of ERP findings by Román et al. (2021). Román et al. tested sentences in which the critical noun phrase was temporarily ambiguous between a direct object and the subject of a sentential complement (*The weary traveler found his suitcase (had been opened)*). Nouns that were implausible objects of the preceding verb (*The weary traveler explained his suitcase…*) yielded no plausibility effects at the noun itself even in the case that the verb was biased towards a direct-object continuation. All the above findings can be accounted for in a constraint-based approach to processing syntactic ambiguities (Garnsey et al. 1997; MacDonald et al. 1994). In this approach, multiple interpretations of the ambiguity are activated in parallel; the relative strength of the activation depends on how much each reading is supported by the information available. In the case of *and*-coordination ambiguities, frequency may bias

towards an NP-coordination. However, the implausibility of the noun and preceding verb may bias towards an S-coordination. This competition may either result in a delay in fully integrating the critical noun, or in a net-result of no visible plausibility effects at the noun (Hoeks et al. 2006; Román et al. 2021). When the critical noun is plausible given the preceding verb, an NP-coordination is favored, resulting in the garden-path effect observed in prior studies when a following verb disambiguates towards an S-coordination (Frazier 1987; Hoeks et al. 2006; Kaan and Swaab 2003; Kaan et al. 2019; Staub and Clifton 2006). The implausible noun in block 1 may therefore have biased the readers' expectation more towards an S-coordination. This bias may have been strengthened by the use of a modifier phrase after the critical noun (*movie about the superhero*). This modification made the noun phrases on either side of *and* less parallel, which may have further biased against an NP-coordination (Frazier et al. 1984).

Our readers may therefore have expected an S-coordination for the implausible conditions, even at the start of block 1. This accounts for the plausibility effect at the sentence-final word in block 1. At the end of the sentence in block 1, it became clear to the readers that there is no verb to salvage the critical noun in the implausible condition. This led to a difference between the plausibility conditions at the end of the sentence: the ERPs in the plausible condition were more positive (typical end-of-sentence effect) than in the implausible condition.

Regarding our main question of how and when readers adapt: we did find evidence of changes in processing as a function of exposure. Over the course of block 1, the end-of-sentence plausibility effect decreased in size. One account of this decrease is that our readers were gradually getting used to the global implausibility and were therefore no longer processing the end-of-sentence anomaly. However, prior studies on adaptation to semantic unexpectancies report that N400 effects are rather robust and not affected by the number of anomalous sentences encountered (Yano et al. 2020). A more likely interpretation of the decrease in the end-of-sentence plausibility effect is that readers started shifting their processing expectations and no longer expected the sentence to continue with a verb in the implausible condition, leading to an increasingly smaller response to the sentence being syntactically incomplete in the implausible conditions.

The end-of-sentence plausibility effect in block 1 changed rather quickly. This contrasts with the observations at the critical noun. The plausibility effect at the noun was not statistically different among the three blocks, in spite of the *and*-coordination being resolved towards an NP-coordination in blocks 1 and 3, and towards an S-coordination in block 2. This suggests that readers were not very sensitive to the immediate changes in distributions of the structures in adapting their processing expectations before the point of disambiguation. The numerical patterns suggest that readers did change their expectation of a particular structure, but that this change occurred rather slowly. The N400 effect at the critical noun was numerically largest and more robust in block 2. This suggest that readers started to more strongly expect an NP-coordination and to integrate the noun with the preceding verb in block 2 compared to block 1. Apparently, readers adhered to an NP-coordination interpretation in block 2 due to the exposure to NP-coordinations in block 1. This was in spite of the fact that the *and*-coordination was systematically resolved towards an S-coordination in block 2. The numerically larger P600 for the plausible vs. implausible condition when the verb in block 2 disambiguated towards an S-coordination also suggests that an NP-coordination was considered for the plausible conditions in block 2.

Recall that the sentence-final word in block 3 (NP-coordinations) elicited an overall negativity which was larger in block 3 than in block 1. This can be accounted for by assuming that readers adapted to the S-coordinations by the end of block 2 and started to expect the sentence to continue with a verb, regardless of plausibility. The end-of-sentence negativity then reflects the unexpected end of the sentence as a verb was missing. The numerically smaller plausibility effect at the critical noun in block 3 than in block 2 supports the interpretation that an NP-coordination was no longer as strongly considered in block 3 as in block 2, inspired by the abundance of NP-coordinations in block 3.

Our results obviously need to be replicated, but if our interpretation is correct, this implies that adaptation to a particular sentence structure is rather slow; processing does not rapidly change in response to just a few sentences of a particular type, in contrast to what has been proposed by some investigators (e.g., Delaney-Busch et al. 2019; Fine et al. 2013). In addition, different aspects of processing may change at different time scales. Changes in the end-of-sentence effect were already noticeable over the course of block 1. We interpreted this effect as readers either getting used to the implausibility or getting used to the sentence missing a verb. The implausibility effect at the critical noun, which we interpreted as a preference towards an NP-coordination was most robust in block 2—although changes in this effect between or within blocks were not statistically significant. Speculatively then, our data suggest that changes in processing of errors (due to something being incorrect or missing) occur rather quickly, whereas changing the processing preference to an alternative structure occurs more slowly.

4.2. *Implications and Limitations*

Important questions in research on adaptation in sentence processing are what readers/listeners adapt to, when, and on what basis? The present study provides a few speculative answers. First, readers quickly get used to sentences being incomplete or anomalous, as witnessed by the change in end-of-the sentence plausibility effects in block 1. Second, these changes occur more quickly than changes in the preference for, and expectation of, a particular structure before the point of disambiguation. If readers quickly adapted their processing preference, a larger N400 plausibility effect would have been seen at the critical word in the NP-coordinations in block 1 compared with the S-coordinations in block 2. Instead, we found a numerically larger plausibility effect at the critical word in block 2. Our results, if replicated, would be problematic for the idea that processing closely reflects the distribution of sentence types in the immediate context.

We acknowledge that the present study is limited and that one needs to be careful in generalizing the results. First, our effects are very small, and differences in the plausibility response at the critical noun between the blocks were not statistically significant. Our sample size may have been insufficient to detect some (interaction) effects and other changes over the course of the study. A replication with a larger group is necessary.

Second, we cannot exclude that changes within and across blocks were affected by changes in attention, task adaptation or other non-linguistic factors, rather than the difference in sentence types between the blocks. To control for this, a second group of participants is needed that would be tested, e.g., only on NP-coordinations, or in a paradigm in which block 1 consists of filler sentences.

Finally, our approach to individual differences was only descriptive. In order to systematically study such effects, a much larger sample is needed, and other measures (e.g., reading exposure, cognitive control, language use and history) need to be collected to understand why certain individuals show a certain type of response (e.g., Grey 2022).

5. Conclusions

Language processing is dynamic: readers and listeners change their processing to accommodate changes and variations in their linguistic environment. The results from the current study suggest that changes in processing in response to exposure to syntactically incomplete sentences or anomalous sentences occur more quickly than changes in structure preferences. In addition, our descriptive data suggest that individuals differ in adaptation. These observations are problematic for the idea that adaptation is an immediate reflection of the properties of the recent context.

Notes

[1] The high pass cutoff of 0.1 Hz differs from what was preregistered (0.01 Hz). We however decided to use the former based on Tanner et al. (2015).

[2] The results from analyses based on individual blocks are discussed in the exploratory analyses in Section 3.4.1 and Supplementary Tables S6–S13. The N400 plausibility effect for the critical noun was significant only in block 2, confirming the analyses reported in the main text. Per a reviewer's suggestion, we also directly compared the plausibility effect on block 1 vs. 2, and block 2 vs. 3 using subsets of the data, and using deviation coding (plausible coded as −0.5, implausible as 0.5). The results are given in Supplementary Tables S3 and S4. The interactions between plausibility and block were not significant (block 1 (coded as −0.5) vs. 2 (0.5): *estimate* = −0.39, *95% CI* [−1.26,0.48]; *SE* = 0.44, *T* = −0.88, *p* = 0.38; block 2 (coded as −0.5) vs. 3 (0.5): *estimate* = 0.60, *95% CI* [−0.28,1.48], *SE* = 0.45, *T* = 1.33, *p* = 0.18).

[3] Based on visual inspection, the preregistered 500–900 msec interval may have been too wide to capture the effect of interest. An exploratory analysis on the average amplitude between 600–800 msec using the preregistered P600 electrode sites yielded a weak effect of plausibility (*estimate* = −0.57, *95% CI* [−1.23, 0.08], *T* = −1.73, *p* = 0.08), see Supplementary Table S5.

References

(Atkinson 2016) Atkinson, Emily. 2016. Active Dependency Completion in Adults and Children: Representations and Adaptation. Ph.D. thesis, Johns Hopkins University, Baltimore, MD, USA.

(Bates et al. 2015a) Bates, Douglas, Reinhold Kliegl, Shravan Vasishth, and R. Harald Baayen. 2015a. Parsimonious Mixed Models. *arXiv* arXiv:1506.04967v2.

(Bates et al. 2015b) Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015b. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48. https://doi.org/10.18637/jss.v067.i01.

(Bentin et al. 1985) Bentin, Shlomo, Gregory McCarthy, and Charles C. Wood. 1985. Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology* 60: 343–55. https://doi.org/10.1016/0013-4694(85)90008-2.

(BNC Consortium 2007) BNC Consortium. 2007. *British National Corpus*. Oxford: University of Oxford, Bodleian Libraries.

(Boudewyn et al. 2019) Boudewyn, Megan A., Adam R. Blalock, Debra L. Long, and Tamara Y. Swaab. 2019. Adaptation to Animacy Violations during Listening Comprehension. *Cognitive, Affective, & Behavioral Neuroscience* 19: 1247–58. https://doi.org/10.3758/s13415-019-00735-x.

(Brothers et al. 2019) Brothers, Trevor, Shruti Dave, Liv J. Hoversten, Matthew J. Traxler, and Tamara Y. Swaab. 2019. Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes. *Neuropsychologia* 135: 107225. https://doi.org/10.1016/j.neuropsychologia.2019.107225.

(Chang et al. 2006) Chang, Franklin, Gary S. Dell, and Kathryn Bock. 2006. Becoming syntactic. *Psychological Review* 113: 234–72. https://doi.org/10.1037/0033-295x.113.2.234.

(Chang et al. 2012) Chang, Franklin, Marius Janciauskas, and Hartmut Fitz. 2012. Language adaptation and learning: Getting explicit about implicit learning. *Language & Linguistics Compass* 6: 259–78. https://doi.org/10.1002/lnc3.337.

(Coulson et al. 1998) Coulson, Seana, Jonathan W. King, and Marta Kutas. 1998. Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language & Cognitive Processes* 13: 21–58.

(De Vincenzi et al. 2003) De Vincenzi, Marica, Remo Job, Rosalia Di Matteo, Alessandro Angrilli, Barbara Penolazzi, Laura Ciccarelli, and Francesco Vespignani. 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85: 280–96. https://doi.org/10.1016/S0093-934X(03)00055-5.

(Delaney-Busch et al. 2019) Delaney-Busch, Nathaniel, Emily Morgan, Ellen Lau, and Gina R. Kuperberg. 2019. Neural evidence for Bayesian trial-by-trial adaptation on the N400 during semantic priming. *Cognition* 187: 10–20. https://doi.org/10.1016/j.cognition.2019.01.001.

(Delorme and Makeig 2004) Delorme, Arnaud, and Scott Makeig. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

(Dempsey et al. 2020) Dempsey, Jack, Qiawen Liu, and Kiel Christianson. 2020. Convergent probabilistic cues do not trigger syntactic adaptation: Evidence from self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 1906–1921. https://doi.org/10.1037/xlm0000881.

(Dempsey et al. 2023) Dempsey, Jack, Qiawen Liu, and Kiel Christianson. 2023. Syntactic adaptation leads to updated knowledge for local structural frequencies. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218231172908.

(Engelhardt and Ferreira 2010) Engelhardt, Paul E., and Fernanda Ferreira. 2010. Processing Coordination Ambiguity. *Language and Speech* 53: 494–509. https://doi.org/10.1177/0023830910372499.

(Fine et al. 2010) Fine, Alex B., Ting Qian, T. Florian Jaeger, and Robert A. Jacobs. 2010. Is there syntactic adaptation in language comprehension? Paper presented at the 2010 Workshop on Cognitive Modeling and Computational Linguistics, Uppsala, Sweden, July 15. Uppsala: Uppsala University, pp. 18–26.

(Fine et al. 2013) Fine, Alex B., T. Florian Jaeger, Thomas A. Farmer, and Ting Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS ONE* 8: e77661. https://doi.org/10.1371/journal.pone.0077661.

(Frazier et al. 1984) Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton, and Kate Ehrlich. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition* 12: 421–30. https://doi.org/10.3758/BF03198303.

(Frazier 1987) Frazier, Lyn. 1987. Sentence processing: A tutorial review. In *Attention and Performance 12: The Psychology of Reading*. Hillsdale: Lawrence Erlbaum Associates, Inc., pp. 559–86.

(Friederici 1995) Friederici, Angela. D. 1995. The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. *Brain Lang* 50: 259–81. https://doi.org/10.1006/brln.1995.1048.

(Garnsey et al. 1997) Garnsey, Susan M., Neal J. Pearlmutter, Elizabeth Myers, and Melanie A. Lotocky. 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 37: 58–93. https://doi.org/10.1006/jmla.1997.2512.

(Grey 2022) Grey, Sarah. 2022. Variability in native and nonnative language: An ERP study of semantic and grammar processing. *Studies in Second Language Acquisition* 45: 137–66. https://doi.org/10.1017/S0272263122000055.

(Guasch et al. 2013) Guasch, Marc, Roger Boada, Pilar Ferré, and Rosa Sánchez-Casas. 2013. NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods* 45: 765–71.

(Gunter et al. 1997) Gunter, Thomas C., Laurie A. Stowe, and Gijsbertus Mulder. 1997. When syntax meets semantics. *Psychophysiology* 34: 660–76. https://doi.org/10.1111/j.1469-8986.1997.tb02142.x.

(Hahne and Friederici 1999) Hahne, Anja, and Angela D. Friederici. 1999. Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 11: 194.

(Harrington Stack et al. 2018) Harrington, Stack, Caoimhe M., Ariel N. James, and Duane G. Watson. 2018. A failure to replicate rapid syntactic adaptation in comprehension. *Memory & Cognition* 46: 864–77. https://doi.org/10.3758/s13421-018-0808-6.

(Hoeks et al. 2006) Hoeks, John C. J., Petra Hendriks, Wietske Vonk, Colin M. Brown, and Peter Hagoort. 2006. Processing the noun phrase versus sentence coordination ambiguity: Thematic information does not completely eliminate processing difficulty. *Quarterly Journal of Experimental Psychology* 59: 1581–99. https://doi.org/10.1080/17470210500268982.

(Jaeger and Snider 2013) Jaeger, T. Florian, and Neal E. Snider. 2013. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* 127: 57–83. https://doi.org/10.1016/j.cognition.2012.10.013.

(Jung et al. 2000) Jung, Tzyy-Ping, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J. Mckeown, Vicente Iragui, and Terrence J. Sejnowski. 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37: 163–78.

(Kaan and Swaab 2003) Kaan, Edith, and Tamara Y. Swaab. 2003. Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research* 17: 621–35. https://doi.org/10.1016/S0926-6410(03)00175-7.

(Kaan and Chun 2018) Kaan, Edith, and Eunjin Chun. 2018. Syntactic Adaptation. In *Psychology of Learning and Motivation*. Edited by Kara D. Federmeier and Duane G. Watson. Cambridge: Academic Press, pp. 85–116.

(Kaan et al. 2019) Kaan, Edith, Corinne Futch, Raquel Fernández Fuertes, Sonja Mujcinovic, and Esther Álvarez de la Fuente. 2019. Adaptation to syntactic structures in native and non-native sentence comprehension. *Applied Psycholinguistics* 40: 3–27. https://doi.org/10.1017/S0142716418000437.

(Kaan et al. forthcoming) Kaan, Edith, Haoyun Dai, and Xiaodong Xu. Forthcoming. Adaptation in L2 sentence processing: An EEG study. *Second Language Research*. https://doi.org/10.1177/02676583231192169.

(Kaschak and Glenberg 2004) Kaschak, Michael P., and Arthur M. Glenberg. 2004. This construction needs learned. *Journal of Experimental Psychology: General* 133: 450–67. https://doi.org/10.1037/0096-3445.133.3.450.

((Kuperberg and Jaeger 2016) Kuperberg, Gina R., and T. Florian Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31: 32-59. 10.1080/23273798.2015.1102299.

(Kuperberg et al. 2020) Kuperberg, Gina R., Trevor Brothers, and Edward W. Wlotko. 2020. A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience* 32: 12–35. https://doi.org/10.1162/jocn_a_01465.

(Kutas and Hillyard 1980) Kutas, Marta, and Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207: 203–06.

(Kutas and Hillyard 1983) Kutas, Marta, and Steven A. Hillyard. 1983. Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition* 11: 539–50. https://doi.org/10.3758/BF03196991.

(Kutas and Hillyard 1984) Kutas, Marta, and Steven A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307: 161–63. https://doi.org/10.1038/307161a0.

(Kuznetsova et al. 2017) Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. https://doi.org/10.18637/jss.v082.i13.

(Lau et al. 2012) Lau, Ellen F., Phillip J. Holcomb, and Gina R. Kuperberg. 2012. Dissociating N400 Effects of Prediction from Association in Single-word Contexts. *Journal of Cognitive Neuroscience* 25: 484–502. https://doi.org/10.1162/jocn_a_00328.

(Lopez-Calderon and Luck 2014) Lopez-Calderon, Javier, and Steven J. Luck. 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci* 8: 213. https://doi.org/10.3389/fnhum.2014.00213.

(MacDonald et al. 1994) MacDonald, Maryellen C., Neal J. Pearlmuttter, and Mark S. Seidenberg. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676–703.

((Metzing and Brennan 2003) Metzing, Charles, and Susan E. Brennan. 2003. When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language* 49: 201-13. https://doi.org/10.1016/S0749-596X(03)00028-7.

(Ness and Meltzer-Asscher 2021) Ness, Tal, and Aya Meltzer-Asscher. 2021. Rational adaptation in lexical prediction: The influence of prediction strength. *Frontiers in Psychology* 12, 622873. https://doi.org/10.3389/fpsyg.2021.622873.

(Nieuwland 2021a) Nieuwland, Mante S. 2021a. How 'rational' is semantic prediction? A critique and re-analysis of Delaney-Busch, Morgan, Lau, and Kuperberg (2019). *Cognition* 215: 104848. https://doi.org/10.1016/j.cognition.2021.104848.

(Nieuwland 2021b) Nieuwland, Mante S. 2021b. Commentary: Rational Adaptation in Lexical Prediction: The Influence of Prediction Strength. *Frontiers in Psychology* 12: 735849. https://doi.org/10.3389/fpsyg.2021.735849.

(Osterhout and Holcomb 1992) Osterhout, Lee, and Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31: 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z.

(Osterhout and Holcomb 1993) Osterhout, Lee, and Phillip J. Holcomb. 1993. Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes* 8: 413–37.

(Osterhout and Nicol 1999) Osterhout, Lee, and Janet Nicol. 1999. On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes* 14: 283–317. https://doi.org/10.1080/016909699386310.

(Pélissier 2020) Pélissier, Maud. 2020. Comparing ERPs between native speakers and second language learners: Dealing with individual variability. In *Interpreting Language-Learning Data*. Edited by Amanda Edmonds, Pascale Leclercq and Aarnes Gudmestad. Berlin: Language Science Press, pp. 39–69.

(Prasad and Linzen 2019) Prasad, Grusha, and Tal Linzen. 2019. Do self-paced reading studies provide evidence for rapid syntactic adaptation? *PsyArXiv*. https://doi.org/10.31234/osf. io/9ptg4.

(Prasad and Linzen 2021) Prasad, Grusha, and Tal Linzen. 2021. Rapid syntactic adaptation in self-paced reading: Detectable, but only with many participants. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47: 1156–72. https://doi.org/10.1037/xlm0001046.

(R Core Team 2021) R Core Team 2021. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

(Román et al. 2021) Román, Patricia, Edith Kaan, and Paola E. Dussias. 2021. Access to verb bias and plausibility information during syntactic processing in adult Spanish–English bilinguals. *Bilingualism: Language and Cognition* 25: 417–29. https://doi.org/10.1017/S1366728921000924.

(Staub and Clifton 2006) Staub, Adrian, and Charles Clifton, Jr. 2006. Syntactic prediction in language comprehension: Evidence from either…or. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32: 425–36. https://doi.org/10.1037/0278-7393.32.2.425.

(Stowe et al. 2018) Stowe, Laurie A., Edith Kaan, Laura Sabourin, and Ryan C. Taylor. 2018. The sentence wrap-up dogma. *Cognition* 176: 232–47. https://doi.org/10.1016/j.cognition.2018.03.011.

(Tanner and Van Hell 2014) Tanner, Darren, and Janet G. Van Hell. 2014. ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia* 56: 289–301. https://doi.org/10.1016/j.neuropsychologia.2014.02.002.

(Tanner et al. 2015) Tanner, Darren, Kara Morgan-Short, and Steven J. Luck. 2015. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology* 52: 997–1009. https://doi.org/10.1111/psyp.12437.

(van Casteren and Davis 2007) van Casteren, Maarten, and Matthew H. Davis. 2007. Match: A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods* 39: 973–78. https://doi.org/10.3758/BF03192992.

(Van Petten and Kutas 1991) Van Petten, Cyma, and Marta Kutas. 1991. Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition* 19: 95–112. https://doi.org/10.3758/BF03198500.

(Van Petten and Luka 2012) Van Petten, Cyma, and Barbara J. Luka. 2012. Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology* 83: 176–90. https://doi.org/10.1016/j.ijpsycho.2011.09.015.

(Weatherholtz and Jaeger 2016) Weatherholtz, Kodi, and T. Florian Jaeger. 2016. Speech perception and generalization across talkers and accents. *Oxford Research Encyclopedia of Linguistics.* Accessed 29 Aug. 2023. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-95.

(Yan and Jaeger 2020) Yan, Shaorong, and T. Florian Jaeger. 2020. Expectation adaptation during natural reading. *Language, Cognition and Neuroscience* 35: 1394–422. https://doi.org/10.1080/23273798.2020.1784447.

(Yano et al. 2020) Yano, Masataka, Shugo Suwazono, Hiroshi Arao, Daichi Yasunaga, and Hiroaki Oishi. 2020. Selective adaptation in sentence comprehension: Evidence from event-related brain potentials. *Quarterly Journal of Experimental Psychology* 74: 645–68. https://doi.org/10.1177/1747021820984623.

(Zhang et al. 2019) Zhang, Wenjia, Wing-Yee Chow, Bo Liang, and Suiping Wang. 2019. Robust effects of predictability across experimental contexts: Evidence from event-related potentials. *Neuropsychologia* 134: 107229. https://doi.org/10.1016/j.neuropsychologia.2019.107229.

Supplementary Materials

Table S1: Analysis of the N400 effect at the critical word, with block 1 as reference level (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.79 ***	-2.35,-1.23	0.29	-6.24	<0.001
Plausibility	-0.30	-0.92,0.31	0.31	-0.96	0.337
Third2	0.52	-0.09,1.13	0.31	1.66	0.097
Third3	0.21	-0.40,0.83	0.31	0.67	0.502
Plaus:Third2	-0.38	-1.25,0.48	0.44	-0.87	0.386
Plaus:Third3	0.22	-0.65,1.09	0.44	0.49	0.625
Random Effects					
σ^2	58.55				
τ_{00} ItemNr	0.88				
τ_{00} Subject	1.07				
N Subject	36				
N ItemNr	216				
Observations	7144				
Marginal R^2 / Conditional R^2	0.001 / 0.033				

* p<0.05   ** p<0.01   *** p<0.001

Table S2: Analysis of the N400 effect at the critical word, with block 3 as reference level (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.58 ***	-2.15,-1.02	0.29	-5.48	<0.001
Plausibility	-0.08	-0.70,0.53	0.32	-0.27	0.790
Third1	0.31	-0.31,0.92	0.31	0.98	0.325
Third2	-0.21	-0.83,0.40	0.31	-0.67	0.502

Plaus:Third1	-0.60	-1.47,0.27	0.44	-1.35	0.176
Plaus:Third2	-0.22	-1.09,0.65	0.44	-0.49	0.625

Random Effects

σ^2	58.55
τ_{00} ItemNr	0.88
τ_{00} Subject	1.07
N Subject	36
N ItemNr	216
Observations	7144
Marginal R^2 / Conditional R^2	0.001 / 0.033

* p<0.05   ** p<0.01   *** p<0.001

Table S3: Exploratory analysis of the N400 effect at the critical word in blocks 1 and 2 using deviation coding (plausible -0.5; implausible 0.5; block 1:-0.5; block 2: 0.5).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.78 ***	-2.21,-1.36	0.22	-8.22	<0.001
Plausibility	-0.50 *	-0.95,-0.04	0.23	-2.14	0.032
Block	0.32	-0.11,0.76	0.22	1.47	0.143
Plausibility:block	-0.39	-1.26,0.48	0.44	-0.88	0.377

Random Effects

σ^2	58.52
τ_{00} ItemNr	0.89
τ_{00} Subject	1.10
τ_{11} Subject.PlausImpl_pl	0.18
ϱ_{01} Subject	0.03
N Subject	36
N ItemNr	216

Observations	4781
Marginal R² / Conditional R²	0.002 / 0.035

Table S4: Exploratory analysis of the N400 effect at the critical word in blocks 2 and 3 using deviation coding (plausible -0.5; implausible 0.5; block 2:-0.5; block 3: 0.5).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.62 ***	-2.03,-1.22	0.21	-7.85	<0.001
Plausibility	-0.38	-0.85,0.09	0.24	-1.58	0.115
Block	-0.01	-0.45,0.43	0.23	-0.04	0.968
Plausibility:block	0.60	-0.28,1.48	0.45	1.33	0.183

Random Effects

σ^2	60.38
τ_{00} ItemNr	0.94
τ_{00} Subject	0.92
τ_{11} Subject.PlausImpl_pl	0.24
ϱ_{01} Subject	0.66
N Subject	36
N ItemNr	216
Observations	4760
Marginal R² / Conditional R²	0.001 / 0.032

Table S5: Exploratory analysis of the P600 effect at the disambiguating verb in block 2 using a narrower time window (600-800 msec, central-parietal electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	2.55 ***	1.80,3.29	0.38	6.68	<0.001
Plausibility	-0.57	-1.23,0.08	0.33	-1.73	0.084
TrialPosition	-0.01	-0.04,0.02	0.01	-0.82	0.413
Plausibility:TrialPosition	0.01	-0.03,0.04	0.02	0.44	0.661

Random Effects

σ^2	64.43
$\tau_{00 \ ItemNr}$	3.88
$\tau_{00 \ Subject}$	2.60
$N_{Subject}$	36
N_{ItemNr}	216
Observations	2337
Marginal R^2 / Conditional R^2	0.002 / 0.093

* p<0.05   ** p<0.01   *** p<0.001

Table S6: Analysis of the N400 effect as a function of trial position at the critical noun in block 1 (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.79 ***	-2.37,-1.22	0.29	-6.16	<0.001
Plausibility	-0.29	-0.89,0.30	0.30	-0.96	0.336
TrialPosition	0.02	-0.00,0.04	0.01	1.61	0.107
Plaus:TrialPosition	-0.01	-0.04,0.02	0.02	-0.49	0.621

Random Effects

σ^2	54.82
τ_{00} ItemNr	0.68
τ_{00} Subject	1.30
N Subject	36
N ItemNr	216
Observations	2384
Marginal R^2 / Conditional R^2	0.002 / 0.037

* p<0.05   ** p<0.01   *** p<0.001

Table S7: Analysis of the P600 effect as a function of trial position at the critical noun in block 1 (500-900 msec, central-parietal electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	0.52	-0.02,1.05	0.27	1.89	0.060
Plausibility	0.08	-0.51,0.67	0.30	0.26	0.798
TrialPosition	0.03 **	0.01,0.05	0.01	2.63	0.009
Plaus:TrialPosition	-0.02	-0.05,0.01	0.02	-1.50	0.134

Random Effects

σ^2	53.74
τ_{00} ItemNr	0.77
τ_{00} Subject	0.96
N Subject	36
N ItemNr	216
Observations	2384
Marginal R^2 / Conditional R^2	0.003 / 0.034

* p<0.05   ** p<0.01   *** p<0.001

Table S8: Analysis of the N400 effect as a function of trial position at the critical noun in block 2 (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.28 ***	-1.84,-0.71	0.29	-4.41	<0.001
Plausibility	-0.69 *	-1.32,-0.06	0.32	-2.14	0.032
TrialPosition	0.00	-0.02,0.03	0.01	0.22	0.824
Plaus:TrialPosition	0.00	-0.03,0.03	0.02	0.15	0.885
Random Effects					
σ^2	61.81				
τ_{00} ItemNr	1.57				
τ_{00} Subject	0.90				
N Subject	36				
N ItemNr	216				
Observations	2397				
Marginal R^2 / Conditional R^2	0.002 / 0.040				

*p<0.05   ** p<0.01   *** p<0.001*

Table S9: Analysis of the P600 effect as a function of trial position at the critical noun in block 2 (500-900 msec, central-parietal electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	1.10 ***	0.61,1.58	0.25	4.43	<0.001
Plausibility	-0.29	-0.88,0.31	0.30	-0.95	0.344
TrialPosition	-0.01	-0.03,0.01	0.01	-1.12	0.264
Plaus:TrialPosition	0.02	-0.01,0.05	0.02	1.50	0.134
Random Effects					
σ^2	54.44				
τ_{00} ItemNr	0.63				
τ_{00} Subject	0.46				
N Subject	36				
N ItemNr	216				
Observations	2397				
Marginal R^2 / Conditional R^2	0.001 / 0.021				

* p<0.05   ** p<0.01   *** p<0.001

Table S10: Analysis of the N400 effect as a function of trial position at the critical noun in block 3 (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	-1.58 ***	-2.10,-1.05	0.27	-5.88	<0.001
Plausibility	-0.10	-0.72,0.52	0.32	-0.33	0.744
TrialPosition	0.02 *	0.00,0.05	0.01	2.05	0.040
Plaus:TrialPosition	-0.01	-0.04,0.02	0.02	-0.74	0.459

Random Effects

σ^2	58.99
τ_{00} ItemNr	0.57
τ_{00} Subject	0.69
N Subject	36
N ItemNr	216
Observations	2363
Marginal R^2 / Conditional R^2	0.002 / 0.023

* p<0.05   ** p<0.01   *** p<0.001

Table S11: Analysis of the P600 effect as a function of trial position at the critical noun in block 3 (500-900 msec, central-parietal electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	1.00 ***	0.47,1.53	0.27	3.71	<0.001
Plausibility	-0.32	-0.92,0.28	0.31	-1.06	0.291
TrialPosition	0.01	-0.01,0.03	0.01	0.64	0.525
Plaus:TrialPosition	-0.01	-0.04,0.02	0.02	-0.52	0.600

Random Effects

σ^2	55.26
τ_{00} ItemNr	0.47
τ_{00} Subject	0.86
N Subject	36
N ItemNr	216
Observations	2363
Marginal R^2 / Conditional R^2	0.001 / 0.024

* p<0.05   ** p<0.01   *** p<0.001

Table S12: Analysis of the N400 effect as a function of trial position at the sentence-final word in block 3 (300-500 msec, central electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	1.33 ***	0.59,2.07	0.38	3.53	<0.001
Plausibility	-0.32	-0.98,0.34	0.34	-0.95	0.345
TrialPosition	-0.02	-0.05,0.00	0.01	-1.95	0.051
Plaus:TrialPosition	0.01	-0.02,0.04	0.02	0.62	0.532
Random Effects					
σ^2	66.47				
τ_{00} ItemNr	2.16				
τ_{00} Subject	2.74				
N Subject	36				
N ItemNr	216				
Observations	2354				
Marginal R^2 / Conditional R^2	0.003 / 0.071				

* p<0.05   ** p<0.01   *** p<0.001

Table S13: Analysis of the P600 effect as a function of trial position at the sentence-final word in block 3 (500-900 msec, central-parietal electrodes).

Parameters	Estimate	95% CI	SE	t	p
Intercept	1.33 ***	0.59,2.07	0.38	3.53	<0.001

Plausibility	-0.32	-0.98,0.34	0.34	-0.95	0.345
TrialPosition	-0.02	-0.05,0.00	0.01	-1.95	0.051
Plaus:TrialPosition	0.01	-0.02,0.04	0.02	0.62	0.532

Random Effects

σ^2	66.47
τ_{00} ItemNr	2.16
τ_{00} Subject	2.74
N Subject	36
N ItemNr	216
Observations	2354
Marginal R^2 / Conditional R^2	0.003 / 0.071

* p<0.05   ** p<0.01   *** p<0.001