Manuscrit accepté pour publication, juillet 2013, version pré-publication

ERPs and Task Effects in the Auditory Processing of Gender Agreement and Semantics in French

Phaedra Royle[1,2], John E. Drury[3], Karsten Steinhauer[2,4]

[1]*École d'orthophonie et d'audiologie, Université de Montréal, Canada,*
[2]*Centre for Research on Language Mind and Brain, Montreal, Canada*
[3]*Department of Linguistics, Stony Brook University, Stony Brook, NY, USA*
[4] *School of Communication Sciences and Disorders, McGill University, Montreal, Canada*

Author e-mails
phaedra.royle@umontreal.ca
john.drury@stonybrook.edu
karsten.steinhauer@mcgill.ca

Corresponding authors

1. **Phaedra Royle**
École d'orthophonie et d'audiologie
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montréal, Qc, Canada, H3C 3J7
514-343-6111 (0925)
phaedra.royle@umontreal.ca

2. **Karsten Steinhauer**
School of Communication Sciences and Disorders,
McGill University,
1266 Pine Avenue West, Montreal,
Quebec, H3G 1A8, Canada
514-398-2413
karsten.steinhauer@mcgill.ca

Work carried out at:
Neurocognition of Language Lab
School of Communication Sciences and Disorders
McGill University
1266 Pine Avenue West
Montreal, Quebec H3G 1A8
Canada

37 PAGES
3 FIGURES / 3 TABLES
WORD COUNT: 11 606 words
Short title: TASK EFFECTS

Abstract:

We investigated task effects on violation ERP responses to Noun-Adjective gender mismatches and lexical/conceptual semantic mismatches in a combined auditory/visual paradigm in French. Participants listened to sentences while viewing pictures of objects. This paradigm was designed to investigate language processing in special populations (e.g., children) who may not be able to read or to provide stable behavioral judgment data. Our main goal was to determine how ERP responses to our target violations might differ depending on whether participants performed a judgment task (Task) versus listening for comprehension (No-Task). Characterizing the influence of the presence versus absence of judgment tasks on violation ERP responses allows us to meaningfully interpret data obtained using this paradigm *without* a behavioral task and relate them to judgment-based paradigms in the ERP literature. We replicated previously observed ERP patterns for semantic and gender mismatches, and found that the task especially affected the later P600 component.

Acknowledgments: We thank Nicolas Bourguignon, Alex Gascon, Véronique Lebel, Benoît Jutras, and Daniel Valois for their help and participation in various stages in the development and running of this experiment. Funding for this project came from the SSHRC (Royle et al, 410-2009-0256), the NSERC (Steinhauer, # RGPGP 312835 and # RGPIN 402678-11) as well as the Canada Research Chair program (Steinhauer, CRC/CFI; project # 201876). No funding agency had any controlling impact on this study.

1. Background

We present an event-related brain potentials (ERP) study examining the influence of the presence versus absence of a behavioral judgment task on adult brain responses to both semantic and morphosyntactic violations in French using a combined auditory/visual sentence-picture matching paradigm. Our motivations were twofold. First, our experimental paradigm was ultimately designed to study language processing and acquisition in young children, who are generally less reliable than adults when asked to perform direct grammaticality or acceptability judgments on linguistic stimuli. Since our experiments with young children using this paradigm *did not* involve such a judgment task, we wanted to establish precisely what kinds of effects the presence or absence of judgment tasks would have on ERP components of interest. This is clearly important for relating child data to adult brain response patterns using our paradigm *with* a judgment task, but also in connecting our findings to the broader literature. Second, the reason such demonstrations are required is that there are fairly mixed results in the literature regarding the task effects on cognitive and linguistic ERPs, as we discuss below. Furthermore, most studies have focused on written language experiments, excepting those done in German. Thus, our study provides valuable data on task effects in psycholinguistic ERP research in French that may also help other researchers decide which paradigms to use in similar research with both adults and children.

ERPs are characteristic patterns of voltage change that are extracted from scalp-recorded electrical brain activity (i.e., EEG signals) time-locked to the presentation of stimuli (Luck, 2005). Systematic differences between ERPs can be used to tease apart the action of systems in the brain involved in different aspects of language processing. For example, ERPs can differ in terms of whether deflections are relatively negative or positive going, when the curves depart from each other (onset latency), how long deflections last (duration), what their distribution is on the scalp (topography), whether the peaks shift in latency, and so on. ERPs have become an increasingly valuable source of empirical evidence for efforts to chart out the functional organization and temporal dynamics of language processing. Previous research has identified a number of ERP effects (components) relevant to individuating mental operations in linguistic sub-domains such as phoneme discrimination, word segmentation, prosodic and intonational phrasing, morphology, conceptual semantics, syntax, discourse processing, and more. However, past research has also made clear that such ERP effects even within particular processing domains can vary in connection with differences in the probability, predictability and modality of the stimuli (visual versus auditory) and in connection with the presence or absence or even the nature of behavioral tasks that participants are asked to perform (Holcomb, 1988; Hoffman, 1990; Friederici, Steinhauer & Frisch, 1999; Hahne & Friederici, 1999, 2002; Steinhauer, Mecklinger, Friederici & Meyer, 1997; Steinhauer 2003).

Thus, when comparing psycholinguistic ERP profiles (e.g., for morphosyntactic processing) between adults and young children who are neither proficient readers nor able to perform in a grammaticality judgment, most available data from adults cannot be used as a control because any observed group differences may actually be due to differences in tasks or modality, rather than processing differences between the two groups. The best approach, therefore, is to test adults with the exact same stimulus materials and presentation paradigm designed for the child population and, in addition, to also test this material with the acceptability judgment task used in the majority of adult ERP studies. This is precisely what the present study does. Importantly, in addition to providing valuable control data for the children's ERPs (not included in this paper), the present study also sheds light on the generalizability of ERP profiles, both cross-linguistically and with respect to their dependence on task requirements during the EEG experiment. This is an issue investigated in only a handful of previous ERP studies, none of which examined French.

In the following, we will first briefly mention relevant dimensions that have been discussed as contributing to task effects. Then we will review previous ERP research conducted in languages other than French. Last, we will use this background to develop specific hypotheses for our present study.

1.1. Why and how do tasks affect the ERP profile?

One thing to keep in mind when using ERPs to study language is that they primarily reflect the various (partly overlapping) *stages of processing* incoming linguistic information (e.g., letter recognition, lexical access, feature checking mechanisms) *in real time*. There is converging evidence that early processing stages (such as letter recognition) are more input-driven (bottom-up) – and thus more robust and automatic – than later 'context-driven' processes (such as metalinguistic judgments), which are more controlled and subject to strategic 'top-down' influences (Shiffrin & Schneider, 1977; Moors & De Houwer, 2006). As a consequence, early 'exogenous' ERP components during the first 100 or 200 ms after word presentation tend to be largely determined by the physical characteristics of the stimulus (e.g., letter font, size, physical contrasts) and less influenced by task requirements, whereas the opposite pattern holds for later 'endogenous' ERP components (such as the P600, a late positivity around 600 ms often found for syntactic anomalies). However, while virtually all late ERP components after 300 ms *can* be modulated by task requirements, some of them may still reflect processes characteristic for normal ‘every-day’ language processing while others are pure artifacts of the experimental setting. For example, a semantic violation as in *John ate democracy* is likely to be perceived as such (an will elicit a corresponding brain response) in almost all circumstances, unless an experimental instruction directs the reader’s attention away (e.g., in a letter monitoring task). On the other hand, brain responses related to specific categorization of this anomaly (e.g., “was it semantic or syntactic?”) might not reflect ‘normal’ language processing and could, therefore, be viewed as an artifact in the laboratory. Since most psycholinguistic ERP studies employ grammaticality (or acceptability) judgment tasks, ‘real’ effects and artifacts can only be teased apart by systematic tests varying the task requirements.

With this general background, we can now ask whether ERP correlates of lexical/conceptual semantic (henceforth: semantic) and morpho-syntactic processing are affected by task demands. As we will see, the answer is influenced by both the mode of processing (single words versus sentences) and modality (visual versus auditory presentation). Since studies on task effects are rare, and since most of them have been done in the reading modality, we review both reading and auditory experiments. For *visual* word processing, the primacy of word meanings over physical features such as font colors in Stroop tasks seemed to strongly suggest automatic semantic processing (the prevailing view in the 1980s). However, studies in the mid 1990s demonstrated that directing a subject’s attention to visual letter features (e.g., *Does the word contain a ‘p’?*) can successfully block semantic processing (e.g., Chwilla, Brown & Hagoort, 1995). This finding is in line with strategic control over higher cognitive functions (e.g., Moors and De Houwer, 2006). However, unlike for single words, in normal *sentence* processing such ‘instruction-induced’ suppression of semantics is less successful (e.g., Gunter & Friederici, 1999). Moreover, in the *auditory* domain, semantic processing (and priming) cannot be suppressed even if participants are asked to focus on an unrelated movie with subtitles (Relander, Rämä & Kujala, 2008). Thus, the retrieval and integration of meaning seems to be an integral part of the ‘default mode’ (or default ‘goal’) of auditory sentence processing, and most studies fail to find substantial task effects for semantics (see below).

Regarding *morpho-syntactic processing*, certain syntax-first models (e.g., Friederici, 2002) have strongly argued for the mandatory nature of an early word-category based phrase structure generation within the first 300 ms after word onset. This process is assumed to be entirely automatic and task independent (Hahne & Friederici, 1999), such that phrase-structure violations (e.g., *He criticized Max's *of proof ...*) should generally elicit an 'early left-anterior negativity' ERP component between 100-300 ms (ELAN) in both modalities. In this model, morpho-syntactic agreement is processed at a later stage that is viewed as less automatic, with agreement violations eliciting later negativies (LANs between 300-500 ms). However, some have questioned the distinction between syntactic and morpho-syntactic processing (and view ELANs and LANs as basically the same; e.g., Ullman, 2004; Hasting & Kotz, 2008). These authors tend to interpret early and late LANs as reflections of automatic (morpho-)syntactic processes, whereas they characterize subsequent positivities (P600s) as reflections of controlled processes (e.g., structural reanalyses).

However, a factor complicating this issue concerns the reliability of LANs, especially for agreement violations in reading studies. Several studies do not find (or at least do not report) LANs for this type of violation, even in adult native speakers (Osterhout, 1995; Tokowitz & MacWhinney 2005, Stroud, Plesch & Phillips, 2006; Foucart & Frenck-Mestre, 2010, 2012). Instead they report late positivities (P600s) only, i.e., components that otherwise follow the LAN in a later time window (600-900 ms). Several researchers argue that only the P600 is a reliable ERP reflection of morpho-syntactic processing, sometimes characterizing the P600 as an ERP correlate of 'implicit' syntactic processing (Tokowitz & MacWhinney, 2005). However, as discussed below, most authors agree that the rather late P600 is perhaps the best candidate for a brain response related to controlled processing (e.g., reanalysis and sentence repair; Hahne & Friederici, 1999). In fact, some have argued it may not reflect genuine psycholinguistic processes at all but rather be related to the metalinguistic classification of grammatical 'wellformedness' (e.g, Bornkessel-Schlesewsky, Kretzschmar, Tune, Wang, Genç, Philipp, et al. 2011), reviving a debate of the mid 1990s as to whether the P600 may be a member of the 'P300 family' of ERP components (e,g,, Coulson, King & Kutas, 1998 vs. Osterhout & Hagoort, 1999). P300s reflect non-linguistic processes of stimulus classification and working memory updating (Donchin & Coles, 1983) that are strongly related to judgment tasks and can, therefore, be viewed as a 'non-linguistic' ERP artifact. The controversy of whether the P600 is a genuine psycholinguistic component (Osterhout & Hagoort, 1999; Steinhauer et al., 1997) or rather a task-dependent effect (Bornkessel-Schlesewsky et al. 2011; Coulson, King & Kutas, 1998) is ongoing.

These conflicting data may be due to the P600 reflecting various cognitive processes rather than being a monolithic component. Teasing apart various subcomponents of the P600 using principal component analysis, Friederici, Mecklinger, Spencer, Steinhauer and Donchin (2001) argued that the P300 may indeed often contribute to, but not entirely account for, P600 effects in psycholinguistic experiments. However, as with most morpho-syntactic ERP data, their findings came from a reading study, and may not generalize to the auditory domain. To summarize, the status and task-dependency of both LAN and P600 components is still controversial (especially for spoken sentences), whereas there is good reason to expect rather robust N400 effects in auditory sentence processing.

Following this line of thinking, we reasoned (as others have) that manipulating the presence versus absence of a judgment task in our experiment can be expected to differentially affect distinct ERP components, depending on whether they reflect more automatic versus more controlled processing. Thus, data from this investigation was expected to also contribute in a broader way to our understanding of the nature of various different types of language-related ERP response patterns.

A last issue that is often overlooked in the analysis and interpretation of ERPs in auditory studies has more recently been raised by several authors (Hasting & Kotz, 2008; Steinhauer, 2003; Steinhauer & Drury, 2012; van den Brink & Hagoort, 2004) and concerns the ‘time-locking’ of ERP effects to the relevant information. Whereas in reading studies using word-by-word presentation methods, the entire word information becomes available at once and can be analyzed relative to word onset time, spoken words unfold over time, such that different types of information (carried by different morphemes) may only be available at different points in time. Thus, depending on whether (morpho-)syntactic information is encoded in prefixes or suffixes, major semantic information (carried by the stem morpheme) may either become available prior to or later than syntactic information. Therefore, in order to quantify the specific time course of effects, auditory ERP analyses should either time-lock analyses to the relevant morphemes or (at least) take the position and duration of the various morphemes into account. For semantic N400s, auditory studies using short content words (of 1-2 syllables) typically find a longer overall component duration compared to reading studies, but the 300-500 ms time window post word onset (typical for reading studies) very often also works to quantify semantic N400 effects of spoken words. Interestingly, if measured relative to the auditory ‘uniqueness’ point of a spoken word, N400s occur much earlier than 400 ms (often even *prior to* the uniqueness point), suggesting that semantic processing is initiated based on incomplete input (much in line with Marslen-Wilson’s, 1973 ‘shadowing’ tasks). For morpho-syntactic ERP components linked to inflectional morphology, most previous auditory studies simply ignored the position of these morphemes relative to word onset, often resulting in misinterpretations of the actual time course of cognitive processes (for discussion see Hasting & Kotz, 2008; Steinhauer & Drury, 2012).

Our experiment, carried out in French, focused on two dimensions of language processing which can be tracked using ERP methods in both adults and children, namely lexical/conceptual-semantic and gender-agreement processing. To our knowledge, oral language gender processing has never been studied in French using ERPs. Three ERP components have been linked to the types of linguistic processing we focus on: The N400, the LAN and the P600. We first discuss these in turn and follow with a number of hypotheses as to the task effects we expect to find with these components. Our review will largely focus on studies using auditory paradigms (with and without picture stimuli) as we chose to use a bimodal auditory-visual method of stimuli presentation, where images were presented simultaneously with short sentences. An advantage of this paradigm is that it obviates reading processes that might not be robust in children, all the while grabbing their attention. This necessarily reduced the pool of studies reviewed, as auditory presentation of complex linguistic stimuli is not yet the norm in ERP research.

1.2 Semantic processing

The N400 component, first described in Kutas and Hillyard (1980), is a negative going brain wave observed between 300 and 500 ms after stimulus presentation. It can be elicited by semantic expectancy violations (Kutas & Federmeier, 2000), and its distribution tends to be most prominent at central and parietal regions. The N400 is extremely reliable and has been described in many neurolinguistic experiments on lexical access or semantic processing in sentences. This component has recently been observed in bimodal (auditory-visual or visual-visual) lexical semantic violation conditions where an incongruous or unexpected image is presented concurrently with an auditory or written utterance, whether it is in a noun phrase or a sentence context (see e.g., Friedrich & Friederici, 2004; Willems et al., 2008).

Although most auditory studies of sentence processing do not show N400 modulation by task, one has. In an auditory sentence processing study, Hahne and Friederici (2002) investigated semantic and syntactic errors in two judgment tasks. One group judged the sentences on all aspects (semantics and syntax), while the second had to ignore syntactic errors while judging semantics only. In pure semantic violations (*Der Vulkan wurde <u>gegessen</u>* 'the volcano was eaten') participants showed comparable N400s in both groups, while *only* in double violation conditions (grammatical + semantic error such as *Das Türschloß wurde *im <u>gegessen</u>* 'the door lock was in-the eaten') N400s disappeared in the group who were instructed to ignore semantic anomalies.

However, most other studies suggest the N400 arises in connection with automatic lexical-semantic processing in sentence contexts. Fischler, Childers, Achariyapaopan and Perry (1985) observed an N400 when participants were either engaged in a sentence learning task or making truth-value judgments. Balconi and Pozzoli (2005) observe N400s to semantic violations (*La lepre nel campo è *dipinta* 'The hare in the field is *painted') in both auditory and visual presentation, irrespective of instructions (paying attention to errors versus no explicit instruction about the experiment). Other types of experimental designs (lexical decision with priming, for example) do show modulation of N400 effects based on proportion of related pairs within lists (Holcomb, 1988; Steinhauer, Nadeau-Noel, Drury & Royle, 2008), or the types of priming used (e.g., paired-priming versus mediated priming, Silva-Pereyra, Harmony, Villanueva, Fernández, Rodríguez, Galán, et al. 1999), however, these experiments usually deal with single word recognition and do not involve the complex and continuous semantic integration believed to occur in sentence processing (Van Petten & Kutas, 1991). Thus is seems that the N400 can be elicited with or without task in sentence processing, but that it can nevertheless be modulated by task demands in certain situations.

1.3 Gender agreement processing

The *Left anterior negativity* (LAN) has been reported for different types of morpho-syntactic violations. This component typically emerges between 300 and 500 ms after stimulus presentation in the case of verb agreement violations such as *As a turtle grows its shell *<u>grow</u> too* (Kutas & Hillyard, 1983) but can appear earlier, especially in the auditory domain and depending on the timing of the error, i.e., word-initial or word-final (see below: Hasting & Kotz, 2008) and, importantly for us, in the presence of agreement errors. Most ERP studies of agreement eliciting LANs have been pursued in the written modality (Barber & Carreiras 2005; Barber, Salillas & Carreiras 2004; Gunter, Friederici & Schriefers 2000; Molinaro, Barber & Carreiras, 2011; O'Rourke, 2008). Only a few auditory studies have investigated agreement, but most report negativities that are similar to, though sometimes more frontal and more sustained than, those found in the visual domain (see e.g., Friederici, Pfeifer & Hahne, 1993; Balconi & Pozzoli, 2005; Rossi, Gugler, Friederici & Hahne, 2006; Hasting & Kotz, 2008; Morgan-Short, Sanz, Steinhauer & Ullman, 2010).

LANs can be argued to reflect automatic processes. Thus, to the extent that automaticity may confer immunity to task effects, one might expect LANs to be elicited independent of whether or not a judgment task is employed. However, here again the relevant data reveal a mixed picture. For example, Osterhout and Mobley (1995) show the LAN appears in English subject - verb agreement (sentence reading) paradigms *only when* participants are asked to perform a judgment task leading these authors to argue that the LAN can indeed be subject to strategic influences. On the other hand, in O'Rourke (2008) agreement errors on determiners and adjectives in written Spanish structures (*El piano esta *rota/roto*, 'the piano is broken.f/.m') elicited LANs even though participants did not perform a judgment task (comprehension probes were used to maintain attention). Further, in some studies the LAN does not seem to be modulated by proportion changes in the stimuli (Gunter, Stowe & Mulder, 1997; Coulson et al. 1998), indicating that it may not be subject to strategic or "list" effects (see Royle, Drury, Bourguignon & Steinahuer, 2012; Steinhauer & Drury, 2012 for relevant discussion). On the other hand, a more recent auditory study by Hasting and Kotz (2008) shows that in two-word sentences with agreement errors (e.g., *er kegelt* 'he bowls' vs. *er *kegelst* '*he bowl') a frontal negativity arises both in grammaticality judgment and in visual distraction conditions (video monitoring while listening to stimuli). Furthermore, the negativity arises rapidly and is short lived (100-300 ms after error onset) in the judgment group, while in the distraction group it is long lasting (100-800 ms) (Hasting & Kotz, 2008). The authors argue that the shorter time-course of the negativity in the first condition reflects a prototypical LAN component for grammatical errors, while the sustained negativity came as a surprise and was linked to a "processing negativity [...] observed in response to task-relevant stimuli in selective attention paradigms" (p. 1216). In contrast, Steinhauer & Drury (2012) reinterpreted these data as a sustained negativity in *both* task groups, which however was cancelled out by a large positivity after 400 ms (P600) in the judgment group only. In fact, these authors provide evidence that most local (E)LAN effects in auditory syntax studies can be explained in terms of a sustained frontal negativity that is temporarily superimposed by a large P600 (e.g., Hahne & Friederici, 1999, Experiment 1). In the absence of P600s, syntax-related negativities tend to last several hundred milliseconds (e.g., Hahne & Friederici, 1999, Experiment 2).

In contrast to the LAN, the positive-going P600 component is a wave emerging between 500 and 1000 ms after visual stimulus presentation (Osterhout & Holcomb, 1992), and between 300 and 1500 ms in the auditory domain (e.g., Hahne & Friederici, 1999; Osterhout & Holcomb, 1993; Steinhauer, Alter & Friederici, 1999). This component is observed for gender agreement violations (see reviews in Molinaro, 2011, and O'Rourke, 2008; see Foucart & Frenck-Mestre, 2010, 2012 for French), verb agreement violations (Balconi & Pozzoli, 2005), syntactic violations (Gunter et al., 1997; Friederici, Hahne & Mecklinger, 1996; Kutas & Hillyard, 1983) as well as syntactically or logically complex structures that are grammatical but result in processing or integration difficulties (Fischler et al., 1983, 1985; Kolk, Chwilla, van Herten & Oor, 2003). The P600 can also be elicited by semantic anomalies in conjunction with N400s (Hagoort, 2003; Steinhauer, Drury, Walenski, Portner, Ullman, 2010). Because it is found in many different contexts, some have proposed that the P600 component indexes processing load relating to language monitoring, and that it is not specifically dedicated to grammatical syntactic processing (Kolk, et al. 2003; Oor, 2003; Steinhauer & Connolly, 2008; van de Meerendonk, Kolk, Vissers & Chwilla, 2010; Bornkessel-Schlesewsky et al. 2011). It has also been argued that the P600 may reflect controlled cognitive processes. Fischler et al. (1985) showed weaker late positivities in the absence of truth-value judgments as compared their presence. Hahne and Friederici (2002) showed that the P600 could be influenced by experimental instructions. In the experiment discussed above in the LAN section, no P600 was observed in response to syntactic violations, when the subjects are asked to judge phrases on semantics while ignoring syntactic structure. However, a direct comparison of these two tasks does not reveal significant differences. Gunter and Friederici (1999) observed smaller P600s in a letter case detection condition (to word-category violations) or no P600 (for verb inflection errors), in comparison to a grammaticality judgment task. Finally, it has been shown that list manipulation can influence the P600 amplitude, increasing its size when the proportion of ungrammatical structures is lowered (Coulson et al., 1998; Gunter, Stowe & Mulder, 1997; Hahne & Friederici, 1999).

On the other hand, Brown van Berkum and Hagoort (2000) show a P600 without grammaticality judgment during auditory syntactic processing of gender (dis)agreement in Dutch. Similarly, O'Rourke (2008) observed P600s in all but one of her conditions involving gender agreement in a Spanish reading task (e.g., *El piano está *rota* 'the.m piano is broken.f'), none involving grammaticality judgment, while Balconi and Pozzoli (2005) find no differences related to task demands (paying attention to errors versus no explicit instruction about the experiment) on P600s to verb agreement errors (*La porta dell'ufficio *sono/è aperte* 'The door of the office *are/is open'), in both visual and auditory modalities. Furthermore, in the study by Hasting and Kotz (2008) discussed above, the P600 arose to grammaticality judgments (for agreement structures vs. phrase structure violations, e.g. *er kegelst* 'he bowl.V' vs *er *Kegel* 'he bowl.N'), but not in the visual distraction condition. Thus, although a majority of studies seem to support the controlled nature of this component, there is some indication that it can be observed without explicitly focusing attention on stimulus semantics or (morpho-)syntax.

1.3.1 Experimental data on agreement

Recent experimental studies using written sentence stimuli with ERPs have revealed brain activation differences between different types of agreement processes, and a number of hypotheses have been put forward to explain these differences. O'Rourke (2008) and Molinaro et al. (2011) present reviews of number and gender agreement processing studies with noun phrases (a.k.a. determiner phrases or DPs) and sentences, most of which use visual word-by-word presentation. We focus here only the gender data involving Det-N or N-Adj agreement. All eleven studies reviewed showed P600s (except one of O'Rourke's long distance conditions), and ten found early negativities (LANs) or N400s for gender violations. O'Rourke initially links the appearance of the LAN to adjacent or nearly adjacent violations (e.g., Det-N or N-Copula-Adj). However, her own experimental data from Spanish show that LANs emerge irrespective of distance, while the P600 is absent in her long distance structures such as *El piano que compramos ayer esta *rota* 'the piano we bought yesterday is broken.f'. She explains the absence of the P600 for these structures by postulating a "good-enough" approach to syntactic processing such that these are incompletely parsed (Ferreira, Bailey & Ferraro, 2002) and interprets the LAN as an automatic index of morpho-syntactic violations, arguing that N400 effects found in some studies indicate that errors are parsed as being semantically incongruous (e.g., *María$_i$ ha prometido a Pedro$_j$ PRO$_i$ ser *estricto con los alumnos.* 'María$_i$ has promised Pedro PRO$_i$ to be *strict.m with the students', Demestre & García-Albea, 2007). Molinaro et al. (2011) highlight the fact that almost all studies show a biphasic pattern with LANs and early or late P600s to incongruent gender conditions. Molinaro and colleagues do not relate LAN/N400 differences in ERP patterns to distance between the agreeing elements, but rather suggest that the presentation context (e.g., isolated N-Adj pairs versus the same elements presented in a sentence) are the cause of the differences observed, the first eliciting an N400-P600 complex while the second results in a LAN-P600, (see e.g., Barber & Carreiras, 2003, 2005). However, the study by Hasting and Kotz (2008) involving a paradigm with two element sentences such as *er *kegelst* ('he *bowl') or *ein *kegelt* (a *bowls), elicit LAN-P600 components. Molinaro, Vespignani and Job (2008) also suggest that morphological transparency of gender marking (for example, processing of gender information on ambiguous nouns in Italian) could promote a lexical N400 rather than a LAN. Finally, a recent series of studies on visual sentence processing in French and second language acquisition, show late positivities (P600) to gender agreement errors but less consistent (non-significant) early negativities in native speakers (Foucart & Frenck-Mestre, 2010, 2012).

In sum, the appearance of N400 or LAN type negativities in agreement conditions is still a question of debate and might be modulated by the languages and structures used to elicit them. In contrast, P600s are clearly more stable responses to agreement errors, at least in grammaticality judgment tasks.

2. Hypotheses

Our experimental paradigm, introduced below, was designed to elicit N400, LAN, and P600 effects. Based on previous findings, we expected our between-group task manipulation to influence these ERP responses as follows. First, because the N400 is systematically larger in sentence contexts with incongruence in the written-visual or auditory-visual input, and does not seem to be reliably affected by task in sentence presentation contexts, we expect that this component will be elicited in our discordant lexical semantic condition (visual vs. auditory input – see below) irrespective of task.

Second, although LANs have rarely been subject to these types of investigations and there is still debate as to their stability and nature (e.g., will it be local, lasting 300-500 ms or sustained, as in Hasting and Kotz, 2008?), it is unclear whether we will elicit these in our paradigm. However, on the assumption that such effects reflect more automatic underlying processes, we believe that if LANs are elicited by grammatical violations, they will be manifest independent of task. Further, following Molinaro et al. (2010), since N-Adj agreement involves idiosyncratic morphological marking in French (Royle & Valois, 2010), it is quite possible that an N400/P600 complex will be elicited for conditions involving these structures rather than a LAN/P600.

Third, the P600 literature suggests that task modulations of this component are a likely outcome. Thus we expect this component will be significantly reduced in the No-task group in comparison to the Task group. The question is whether the P600 will survive the absence of a goal in the form of a grammaticality judgment. A summary of our hypotheses is presented in Table 1.

Table 1 about here

3. Materials and Methods

3.1 Participants

Twenty-three neurotypical adults aged 18 to 35 participated in the experiment. All were right-handed (as assessed with the Oldfield, 1971), and had French as their mother tongue and their everyday language. All were questioned about their developmental and linguistic history and none had a history of learning disabilities, neurological damage, hearing-loss or drug abuse. They were recruited from the university student populations of Montréal, Québec. Each signed a consent form before participating in the experiment. Participants were randomly assigned to the Task (grammaticality judgment) or No-task (listening for comprehension) condition. All received 45$ for their participation. Following data inspection 15 participants were retained for the study. The eliminated data sets presented high levels of artifacts (blinking or sweating) or electrode reading problems that contaminated the ERP signal. Of the remaining participants, 5 women and 2 men participated in the Task condition, and 5 women and 3 men were in the No-Task group.

3.2 Stimuli

Stimuli selection was constrained by age of acquisition norms, as we developed the stimuli to be used on younger populations after running it on adults. 12 French nouns and 8 adjectives acquired before the age of three on Quebec French norms were used (Trudeau, Frank & Poulin-Dubois 1999). Online corpora were checked to provide frequency norms for selected items. These include the Manulex (Lété, Sprenger-Charolles & Colé 2004), Novlex (Lambert & Chesnet, 2001) and Lexique (New, Pallier, Ferrand & Matos, 2001), all frequency databases for written and spoken French, the first two based on child books for grade-school. Nouns and adjective pairs were chosen based on their imageability. All critical nouns had consonantal onsets (to avoid elision -- vowel erasure in the preceding determiner -- an obligatory phonological process in French) and were unvoiced, to make sentence splicing more efficient. Half of the nouns were feminine and half of the adjectives were invariable (these have the same form in the feminine and the masculine). All stimuli were matched on frequency, age of emergence and length within lexical category (see Table 2 for details). Incongruent nouns in the semantic condition (see Appendix, 1a/b) belonged to a different semantic domain and did not share the initial phoneme of the congruent form (e.g., *camion* [kamiõ] 'truck' – *soulier* [sulje] 'shoe'). Nouns and color adjectives were combined to create 48 feminine and 48 masculine adjective-noun pairs that were inserted into carrying sentences containing a lead in (*je vois* 'I see' or *il y a* 'there is'), and a sentence continuation prepositional phrase (*sur la table* 'on the table' or *dans la boîte* 'in the box') to avoid wrap up effects in the ERPs. A visual stimulus was created for each sentence. A professional artist created drawings emphasizing the relevant properties (colours) of interest. The drawings maintained a constant level of visual complexity, avoiding superfluous or distracting details.

An example is presented in the Appendix. The sentences were presented in the context of a story about an alien (Zilda) coming to visit Québec. She has to practice her French on the way to Earth.

Table 2 about here

Auditory stimuli were recorded at 44.1KHz in a sound attenuated booth using a Sony DAT recorder (PCM-M1 recorder, 1997). All sentences were spoken by a native French Canadian actor who was trained to clearly articulate the words with natural intonation while avoiding co-articulation at word boundaries. All conditions within a given block of stimuli were recorded together. However, only grammatical sentences were recorded, as tests in our lab have shown that ungrammatical structures cause subtle but significant slowing in production as well as intonation modifications, even with trained speakers. Speaker voice intensity was maintained at 65 dB (± 5 dB) throughout the recording session by monitoring her with a sonometer, to reduce post-recording manipulations. Following stimuli recording, the sentences were normalized and spliced using Cool Edit software (Syntrillium Software, Phoenix, AZ). Sentences were normalised at 70%. Grammatical and ungrammatical sentences were constructed by cross-splicing the original recordings to prevent acoustic confounds (ex. … le / camion / brun … – … la / sacoche / brun …. ‘the.m / truck / brown.m – the.f / handbag / brown.m’) cutting at the onset of the determiner, the noun, the adjective, and the prepositional phrase (see example in 1 illustrating splicing points).

1. Je vois | une | sacoche | brun | sur la table
   I see | a/one | handbag | brown | on the table.

On adjectives, the point of recognition (depending on whether the masculine or the feminine form was the stimulus) was at 125/160 ms for *vert/verte* [vɛʁ/vɛʁt] 'green', 150/275 ms for *brun/brune* [bʁỹ/bʁyn] 'brown', 320 ms for gris/grise [gʁi/gʁi:z] 'grey', and 400 ms for *blanc/blanche* [blã/ blãʃ] 'white'. The point of recognition was established by a trained audiologist (V. Lebel) who listened to the items, using gating techniques, until she could establish whether the adjective was feminine or masculine. Note that for the present analyses, time-locked to the onset of the adjective in the agreement condition, the exact position of the recognition point is not essential and is provided for completeness.

Conceptual semantic processing was investigated by creating semantic violations where the image did not correspond to the noun presented in the auditory stimulus (e.g., the sound file describes ‘a white handbag on the table’ and the image is of a white CHAIR on a table). Morphosyntactic agreement processing was studied by creating gender mismatches between the noun and adjective in the auditory stimulus (e.g. *une saccoche blanc/*blanche sur la table*, ‘a handbag white.m/.f on the table'), that is by splicing in a feminine adjective with a masculine noun, or vice versa.

192 stimuli sentences were presented to participants in each list (192 additional items using the same materials were included as items for another experiment not discussed in this paper). Forty eight trials were congruent and grammatical (note that the 48 incongruent semantic trials were also grammatical). Each incongruent condition was associated with a congruent one. As each noun (12) and adjective (8) was combined to create our sentence materials, every participant heard each noun 32 times and each adjective 48 times throughout the entire experiment. Four different pseudo-randomized lists were created for stimuli presentation, taking care to balance different conditions and control items across halves and quarters of the lists. The sentences were presented in the context of an ‘Alien learning paradigm’ where an alien was learning French in preparation for a visit to Quebec, on the way to earth (in the first session) and while unpacking her boxes in her new house (in the second session). A story containing filler sentences, images and animations was interspersed throughout the experiment to maintain interest and attention.

3.3 Procedure

Participants were fitted with an EEG cap and were seated in a comfortable chair in a sound-attenuated and electromagnetically shielded booth, at a distance of ~ 1m from a computer screen. Participants were presented with one of four lists during the recording session. Insert ear plugs were used to present the auditory stimuli, while the images appeared on the screen. The experiment was divided into two 40-minute halves with 20 and 22 blocks of stimuli respectively. A pause was programmed every ten minutes on average. Examples presented at the beginning of the experiment explained that Zilda the alien was learning a new language and could make mistakes. In the Task condition, the participants were instructed to listen to each sentence and judge its appropriateness using a mouse key. In the No-task condition participants were asked to listen to the sentences and to pay attention to the story in order to answer questions that would subsequently be asked by the experimenter. Every five minutes the experimenter would pause the experiment to ask a question on the content of the story

(*How many moons were there?*) or the preceding sentence (*What color was the shoe?*). These did not necessarily focus on the experimental incongruencies and were added after pilots tests showed that participants might not be paying attention and could become distracted. No other instructions were given. The visual stimulus was presented at the onset of the sentence and remained on the monitor until the end of the auditory stimulus. Visual-auditory presentation of this type has been shown to elicit typical N400s for semantic incongruencies (Willems et al, 2008) as well as typical N400/P600 complexes for thematic-role reversals (Wassenaar & Hagoort, 2007). A lag of 1000 ms followed the disappearance of the image, after which a prompt was given to Task group participants to make a grammatical judgment. This was followed by a visual prompt ("- -") that was presented for two seconds to indicate the interval allotted for eye blinking (thereby dramatically reducing the number of eye-blink contaminated trials). Following this a new sentence-picture pair was presented. In the No-task group, the prompt for a decision was eliminated while pauses for questions were added.

3.4 Data analyses

The EEG was recorded continuously with a 500 Hz sampling rate from 64 cap-mounted electrodes (Ag/AgCl Electrocap International Inc., Eaton: OH). Four additional electrodes were attached above and below the left eye as well as on both temples to monitor vertical and horizontal eye movement respectively. All impedances were maintained below 7 kΩ) The EEG was amplified using a Neuroscan SynAmps2 DC amplifier referenced to the right mastoid. All subsequent steps of EEG/ERP data processing and analysis were carried out with the EEProbe software package (ANT; Enschede, The Netherlands). The electrodes covered frontal, central, parietal, temporal and occipital lobes (FP1, FP2, F3, F4, F7, F8, Fz, C3, C4 Cz, P3, P4, Pz, T3, T4, T5 T6, O1, O2, Oz). Offline, data were re-referenced to linked mastoids and filtered with a bandpass of 0.3 to 40Hz. Trials contaminated with eye blinks and other artifacts were rejected using a 30 μV criterion (resulting in a data loss of 9 %, evenly distributed across conditions). All uncontaminated trials were entered into the final analysis. ERP averages were computed in a 1600 ms time epoch, including a 100 ms prestimulus baseline interval (-100 to 1500 ms).

INSERT FIGURE 1 ABOUT HERE

Following a visual inspection of each participant's ERP file, data sets that were considered to be too noisy or containing insufficient data points per condition were excluded (9 cases). In the remaining subjects, 22% of critical trials on average were rejected due to artifacts (eye movements or electrode drifts). An average of 37.3 items per violation condition were analyzed. These did not differ significantly by condition or by participant group (all *p*s > 0.1). Based on our visual inspection of the data, we selected specific time-windows for statistical analyses of the grammatical or semantic errors.

For each time-window, and each condition separate analyses time-locked to the critical stimuli (determiner, noun or adjective) were performed for lateral and central electrodes. The ANOVA for the central electrodes had three factors: GROUP (Task and No-task), COND(ITION) (congruent and incongruent), and ELEC(TRODE) (Fz, Cz, Pz and Oz). For the lateral analysis, the ANOVA included 5 factors: GROUP (Task and No-task), COND (congruent and incongruent), LAT(ERALITY) (2 levels, more and less lateral), HEM(ISPHERE) (right and left) and ANT(ERIORITY/POSTERIORITY) (4 levels). An alpha of $p < 0,05$ was used for all statistical analyses. A Greenhouse-Geisser correction for sphericity was used for analyses with more than one degree of freedom in the numerator.

4. Results

4.1 Semantic violation condition

Semantic mismatches elicited two distinct ERP components on the incongruent noun (see ERP waves comparing groups and conditions in Figure 2). First a centro-parietal negative wave (N400) emerged around 200 ms at both midline and lateral electrodes and lasted until approximately 500 ms after onset. Second, a positive wave (P600) emerged around 650 ms and lasted until 1000 ms post word onset. The late positivity has a smaller amplitude and is restricted to occipital areas in the no-task group, while it has a larger amplitude, is widely distributed over posterior sites and is slightly left lateralized in the task group (see Figure 2, for central electrodes and voltage maps). Statistical analyses on time windows 200 - 500 ms, and 650 - 1000 ms show that in the earliest time-window a main effect of COND was found in both lateral and midline analyses (Lateral, $F(1,13)$ 25.67, $p < 0.001$, η^2 1.97; Midline $F(1,13)$ 24.48, $p < 0.001$, η^2 1.87). No main effects of GROUP were found (all p's > 0.1), and no interactions of GROUP with other factors. The effect of COND showed interactions with LAT in the lateral analysis $F(1,13)$ 39.90, $p < 0.001$, η^2 3.04, and ELEC in the midline analysis, $F(3,39)$ 5.05, $p < 0.05$, η^2 0.38. These interactions highlight the fact that the N400 is distributed over the parietal and occipital regions of the scalp and are stronger in central versus lateral electrodes. Analyses of the later time-window shows no main effect of either GROUP or COND (all p's > 0.1) in central and lateral electrodes. However, a trend towards a GROUP*COND interaction $F(1,13)$ 3,58, $p < 0.1$, η^2 0.28, and significant interactions of COND*LAT, $F(1,13)$ 4.82, $p < 0.05$, η^2 0.37, COND*LAT*GROUP, $F(1,13)$ 4.67, $p < 0.05$, η^2 0.36, COND*ANT, $F(2,26)$ 28.60, $p < 0.001$, η^2 2.2, COND*ANT*HEM, $F(2,26)$ 9.11, $p < 0.01$, η^2 0.71, and COND*ANT*HEM*GROUP, $F(2,26)$ 7.11, $p < 0.01$, η^2 0.55, are found. Only the Task group shows a COND*LAT interaction, $F(1,7)$ 8.29, $p < 0.05$, η^2 1.19, as well as a COND*ANT*HEM interaction, $F(2,14)$ 25.42, $p < 0.001$, η^2 3.69. In the midline analysis, significant interactions of GROUP*COND, $F(1,13)$ 5.60, $p < 0.05$, η^2 0.43, COND*ELEC, $F(3,39)$ 19.10, $p < 0.001$, η^2 1.47, and a trend towards an interaction of COND*ELEC*GROUP, $F(3,39)$ 2.85, $p < 01$, η^2 0.22, are also found. Only the Task group shows a main effect of COND, $F(1,7)$ 7.31, $p < 0.05$, η^2 1.04, while both show an interaction of COND*ELEC (Task; $F(1,7)$ 5.65, η^2 0.81, $p < 0.05$; No-task: $F(1,6)$ 12.59, $p < 0.01$, η^2 2.10). These interactions support our observation that the late positivity is restricted to occipital areas in the No-task group, and is more salient, more widely distributed and slightly left-lateralized in the Task group.

INSERT FIGURE 2 ABOUT HERE

4.2. Adjective-noun agreement violations

Adjective agreement errors elicited two distinct ERP components on the incongruent adjective (see ERP waves and voltage maps in Figure 3). A first negative-going wave with a somewhat left-lateralized distribution (LAN) emerged at approximately 400 ms after adjective onset and lasted until between 600 and 700 ms (depending on the group), while a second positive-going (P600) wave emerged between 550 and 750 ms (depending on group) and lasted up to 1100 ms post onset. The negativity in the No-task group shows a frontal left-lateralized negativity while the Task group shows more bilateral temporal negativities. Furthermore, the relative positivity is much larger and has a more central and left-lateralized distribution in the Task group, while being observable only in posterior central electrodes in the No-task group. See Figure 3. We performed statistical analyses on time windows 450—675 ms, and 675—1100 ms. Analyses for the 450—675 ms time-window reveal main effects of CONDITION in both lateral and midline electrodes (Lateral, $F(1,13)$ 10.39, $p < 0.01$, η^2 0.80; Midline $F(1,13)$ 5.50, $p < 0.05$, η^2 0.42) and no main effects of GROUP nor interactions ($p > 0.1$). In the lateral analyses, additional interactions of COND*LAT*HEM, $F(1,13)$ 9,48, $p <$

0.01, η^2 0.73, COND*ANT*HEM $F(2,26)$ 8.28, $p < 0.01$, η^2 0.64, and COND*LAT*ANT*HEM $F(2,26)$ 5.61, $p < 0.01$, η^2 0.43, were found, in addition to an interaction of COND*LAT*ANT*HEM*GROUP, $F(2,26)$ 3.93, $p < 0.05$, η^2 0.30. This interaction of COND*LAT*ANT*HEM is observed in the No-task group only $F(2,12)$ 6.42, $p < 0.05$. This final interaction supports the observation that the No-task group shows a much more focal and left-lateralized LAN than the Task group. Analyses for the 675—1100 ms time-window reveal a main effect for CONDITION in both analyses (Lateral, $F(1,13)$ 20.13, $p < 0.001$, η^2 1.55; Midline $F(1,13)$ 26.01, $p < 0.01$, η^2 2.0), as well as interactions of GROUP*COND (Lateral, $F(1,7)$ 10.16, $p < 0.01$, η^2 0.78; Midline $F(1,7)$ 11.69, $p < 0.01$, η^2 0.90. Only the Task group shows a main effect of COND (Lateral, $F(1,7)$ 39.21, $p < 0.001$, η^2 5.60; Midline $F(1,7)$ 37.51, $p < 0.001$, η^2 5.36). In the lateral analysis, interactions of COND*LAT, $F(1,13)$ 23.47, $p < 0.001$, η^2 1.81 (and COND*LAT*GROUP, $F(1,13)$ 13.49, $p < 0.01$, η^2 1.04), COND*ANT, $F(1,13)$ 8.53, $p < 0.01$, η^2 0.66, COND*LAT*HEM, $F(1,13)$ 4.99, $p < 0.05$, η^2 0.28, and COND*ANT*HEM, $F(2,26)$ 5.56, $p < 0.05$, η^2 0.43, are found. An analyses of the COND*LAT effects within groups revealed that this interaction was only significant in the Task group, $F(1,7)$ 25.78, $p < 0.01$, η^2 3.68. In the midline analysis interactions of COND*ELEC, $F(3,39)$ 5.18, $p < 0.05$, η^2 0.40 (and COND*ELEC*GROUP, $F(3,39)$ 7.57, $p < 0.001$, η^2 0.58), were observed. Both groups showed this interaction in the within-group analysis (Task: $F(3,21)$ 8.40, $p < 0.05$, η^2 1.20; No-task: $F(3,18)$ 4.31, $p < 0.05$, η^2 0.72). These results support the observation that the relative positivity has a larger amplitude, and that it is more widely distributed in the Task group, while it is of smaller amplitude and more focal in the No-task group: it is in fact, non significant.

INSERT FIGURE 3 ABOUT HERE

5. Discussion

Data from our experiment showed differential modulation of ERP components depending on the task, as we had expected. In the semantic violation condition, the N400 arose in both groups of participants. According to our predictions, in a task such as ours using a visual-auditory presentation paradigm, the N400 should not be modulated (i.e, by grammaticality judgment vs. comprehension). This result confirms our hypothesis that the N400 can be viewed as a rather automatic component of semantic lexical processing, at least in tasks involving processing of concurrent auditory nominal and visual stimuli within sentence contexts, and does not seem to be modulated by attention, when this attention is directed towards grammaticality judgment versus global comprehension. We also observed a positivity following the N400 in this condition. This positivity was stronger in the Task group. As mentioned in our review above, P600s have been observed to follow the N400 in various ERP studies of semantic processing in sentences (see Steinhauer et al., 2010 and references therein) or priming (Holcomb, 1988). We believe that this later positivity reflects attempts to reanalyze or repair the semantic anomaly in the auditory visual mismatch (Steinhauer & Conolly, 2008). This result replicates previous findings of amplitude modulation as a function of task, with the P600 being larger in judgment conditions.

In the Noun-adjective discord condition we observed the expected biphasic LAN-P600 in both groups. The negative-going wave was different from the one elicited in the semantic condition in important ways. Visual inspection revealed strong lateralization of the negativity (bilaterally in the Task group and left-laterally in the No-task group), resembling a typical LAN. The LAN was also more frontal in the No-task group, showing a maximal effect in left-frontal electrodes (F3, F7). Thus the effect of the judgment task on this component was to make it more bi-lateral, not larger in amplitude. Although both groups show a P600 in the later time window, its amplitude is significantly larger and it is more broadly distributed in the Task group (see Figure 3). Because the P600 also occurred in the No-task group, this signals that some aspects of processing indexed by the P600 might reflect genuine psycholinguistic processes (in line with Steinhauer et al., 1997 and Osterhout & Hagoort, 1999). Task however, has an impact on this component, implying that it is at least partly indexing controlled metalinguistic processes. This finding is in line with studies emphasizing the contribution of well-formedness judgments to P300-like P600 sub-components (Coulson et al., 1998; Bornkessel-Schlesewsky et al., 2011), but also confirms previous findings suggesting that P600s are not just P300s (e.g., Friederici et al., 2001).

Table 3 about here

In Table 3 we present a summary of our results for task effects in our different conditions. We had predicted that early and presumably more automatic components should be less influenced by task demands while those that are later and more controlled would not be. These expectations were not always met. The LAN was modulated by TASK, showing more left lateralization in the no-task group in the adjective agreement condition. Congruent with our predictions, the P600 showed significant task effects in both the semantic and adjective mismatch conditions.

It appears that ERP component modulation occurs essentially on late ERP signatures. On the one hand, imposing a grammaticality judgment task will promote P600 effects in terms of its latency and amplitude, while being in the context of a comprehension task still elicits P600s but with reduced latencies and amplitudes. The N400 was unsurprisingly unaffected by task. Our study extends previous findings by showing that the LAN can also be reliably elicited without task requirements. Importantly, the same components appear in both groups, indicating that the use of grammaticality judgment tasks is not necessary to elicit these components. In addition, our study extends previous work on visual-auditory paradigms, none of which, to our knowledge, has tested agreement processing. Similar to the Dutch study by Willems et al. (2008), we observed N400s to visual-auditory mismatches of semantic-lexical information. Similar to a majority of visual (sentence presentation) and auditory studies of gender agreement in other languages (including Dutch, Italian, Spanish and German) we observed biphasic LAN/P600 patterns to agreement errors.

6. Conclusion

Our data converge with the literature showing that Task can modulate ERP components arising in language-based processing experiments. Our data on auditory language processing in French bring a number of new points to bear. First, similar cognitive components arise in auditory ERPs as those found in previous reading experiments for semantic-visual mismatch and gender concord studies. Furthermore, the effects of gender mismatch in French are similar to those found in other languages, despite the highly idiosyncratic nature of gender marking on French adjectives. In addition, the data show that ERPs can be modulated by task, but not as predicted by some. As in all purely conceptual semantic sentence studies we are aware of, the N400 was not modulated by task. We believe that the minimal attention necessary in the No-task group condition mirrors a normal mode of sentence processing and is a useful tool for the study of semantic processing without express orientation of attention in special populations such as children.

Second, the only difference we found on LANs between groups was in their degree of lateralization. The overall effect we found was stronger bi-lateralization during grammaticality judgment, while the more naturalistic comprehension-based auditory condition resulted in frontal left-lateral LANs for agreement errors. However, no group differences were found in terms of amplitude or timing. Variability in lateralization in auditory LANs is a frequent finding, and for now we refrain from interpreting this effect.

Third the P600 was extremely sensitive to task, being much less salient or widespread in the No-task group overall. It appears that this component might be highly linked to orientation of attention to grammar or to well-formedness of the structures. When participants were not focusing on ungrammaticality per se, they showed reduced P600s. Its presence in the No-task group suggests the P600 is not just a task related component but can be interpreted as a linguistically-based component of analysis and repair. This is a promising result, as we can expect to elicit similar components in special populations without task, assuming their processing abilities are similar to neurotypical adults.

In conclusion, it appears that it is indeed possible to elicit typical linguistic ERP components without asking participants to make grammaticality judgments. We have shown this using a paradigm specifically designed for children in terms of visual and auditory stimuli, while taking into account psycholinguistic factors (age of acquisition, structure complexity), that probes both lexical semantics and grammatical processing. This is an interesting point as it allows us to extend the study of these components to populations generally assumed to be unable to make reliable grammaticality judgments. These include children with and without language impairment, for whom it is unclear whether their grammatical processing abilities are abnormal, adult-like or even going through a process of maturation, but also patients with brain lesions. Using ERPs will allow us to probe these questions without asking children to make metalinguistic judgments on stimuli. We are presently addressing the issue of component maturation in French-speaking children aged 5 to 9 years of age (Courteau, Royle, Gascon, Marquis, Drury & Steinhauer, 2013; Royle & Steinhauer, 2013) and children with specific language impairment (SLI).

Bibliography

Balconi, M., & Pozzoli, U. (2005). Comprehending Semantic an Grammatical Violations in Italian. N400 and P600 Comparison with visual and Auditory Stimuli. *Journal of Psycholinguistic Research, 34*(1), 71-98.

Barber, H., & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex, 39*(3), 465-482.

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience, 17*(1), 137-153.

Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. C. Jr. (Eds.), *On-line study of sentence comprehension; eye-tracking, ERP and beyond*. Brighton, UK: Psychology Press.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., et al. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language, 117*, 133-152.

Brown, C. M., Hagoort, P., & Chwilla, D. J. (2000). An Event-Related Brain Potential Analysis of Visual Word Priming Effects. *Brain and Language, 72*, 158-190.

Brown, C. M., van Berkum, J. J. A., & Hagoort, P. (2000). Discourse before gender: An event-related brain potentials study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholinguistic Research, 29*, 53-68.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32*, 274-285.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes, 13*(1), 21-58.

Courteau, E., Royle, P., Gascon, A., Marquis, A., Drury, J. E., & Steinhauer, K. (2013). Gender concord and semantic processing in French children: An auditory ERP study. In S. Baiz, N. Goldman & R. Hawkes (Eds.), *Proceedings of the 37th Annual BUCLD* (Vol. 1, pp. 87-99). Boston: Cascadilla.

Demestre, J., & García-Albea, J. E. (2007). ERP Evidence for the Rapid Assignment of an (Appropriate) Antecedent to PRO. *Cognitive Science, 31*(343-354).

Donchin, E., & Coles, M. G. H. (1988). In the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357-374.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science, 11*, 11-15.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry Jr., N. W. (1983). Brain Potentials Related to Stages of Sentence Verification. *Psychophysiology, 20*(4), 400-409.

Fischler, I., Childers, D. G., Achariyapaopan, T., & Perry Jr, N. W. (1985). Brain potentials during sentence verification: Automatic aspects of comprehension. *Biological Psychology, 21*(2), 83-105.

Foucart, A., & Frenck-Mestre, C. (2010). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*(3), 379-399.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language, 66*(1), 226-248.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78-84.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1219-1248.

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*, 305-323.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*, 183-192.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory and Cognition, 27*(3), 438-453.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-months-olds: processing known words in picture contexts. *Journal of Cognitive Neuroscience, 16*(88), 1465-1477.

Gunter, T. C., & Friederici, A. D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology, 36*, 126-137.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of*

Cognitive Neuroscience, 12, 556-568.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34*, 660-676.

Hagoort, P. (2003). Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience, 15*(6), 883-899.

Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philisophical Transactions of the Royal Society B*(363).

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*(2), 194-205.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*(2), 194-205.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*, 339-356.

Hasting, A. S., & Kotz, S. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cogniive Neuroscience, 20*(7), 1207-1219.

Hoffman, J. E. (1990). Event-Related Potentials and automatic and controlled processes. In J. W. Rohrbaugh, R. Parasuraman & J. R. Johnson (Eds.), *Event-Related Brain Potentials Basic Issues and Application* (pp. 145-157). New York, NY: Oxford University Press.

Holcomb, P. J. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain and Language, 35*(1), 66-85.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85*(1), 1-36.

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

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition, 11*, 539-550.

Lambert, E., & Chesnet, D. (2001). Novlex: une base de données lexicales pour les élèves de primaire. *L'Année Psychologique, 101*, 277-288.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The Role of Structural Prediction in Rapid Syntactic Analysis. *Brain and Language, 98*, 74-88.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX : A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments, & Computers, 36*, 156-166.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 48*(8), 908-930.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228*, 161-176.

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin, 132*(2), 297-315.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*, 154-193.

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*, 522-523.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'Année Psychologique, 101*, 447-462.

Oldfield, R. C. (1971). The Assessment and Analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97-115.

O'Rourke, P. L. (2008). *The Nature of Syntactic Gender Processing in Spanish: an ERP Study.* Unpublished Ph.D., University of Arizona, Tucson, AZ.

Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 14*(1), 1-14.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Proceses, 8*, 413-438.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 24*, 739-773.

Relander, K., Rämä, P., & Kujala, T. (2008). Word Semantics Is Processed Even without Attentional Effort. *Journal of Cognitive Neuroscience, 21*(8), 1511-1522.

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from Event-related potentials. . *Journal of Cognitive Neuroscience, 18*(12), 2030-2048.

Royle, P. (2011). On the Existence of C/Ø Alternations in French Adjectives: Theoretical and Empirical Questions *Proceedings of the 17th ICPhS* (pp. 1730-1733). Hong Kong.

Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: A masked ERP priming study of French verbs. *Neuropsychologia, 50*(14), 3542-3553.

Royle, P., & Steinhauer, K. (2013). Using auditory-visual ERP experiments to study normal and impaired child language development. *Neurophysiologie Clinique/Clinical Neurophysiology, 43*(1), 77-78.

Royle, P., & Valois, D. (2010). Acquisition of adjectives in Quebec French as revealed by elicitation data. *Journal of French Language Studies, 10*(3), 313-338.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*, 127-190.

Silva-Pereyra, J., Harmony, T., Villanueva, G., Fernández, T., Rodríguez, M., Galán, L., Díaz-Comas, L., Bernal, J, Fernández-Bouzas, A., Marosi, E., & Reyes, A. (1999). N400 and lexical decisions: automatic or controlled processing? *Clinical Neuropsychology, 110*, 813-824.

Silva-Pereyra, J., Rivera-Gaxiola, M., Aubert, E., Bosch, J., Galaán, L., & Salazar, A. (2003). N400 during lexical decision tasks: a current source localization study. *Clinical Neurophysiology, 114*, 2469-2486.

Silva-Pereyra, J., Rivera-Gaxiola, M., & Kuhl, P. K. (2005). An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. *Cognitive Brain Research, 23*(2-3), 247-258.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language, 86*(1), 142-164.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*(2), 191-196.

Steinhauer, K., & Connolly, J. F. (2008). Event-related potentials in the study of language. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the Cognitive Neuroscience of Language* (pp. 91-104). New York: Elsevier.

Steinhauer, K., & Drury, J. E. (2011). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language, 120*(2), 135-162.

Steinhauer, K., Drury, J. E., Walenski, M., Portner, P., & Ullman, M. T. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia, 48*(6), 1525-1542.

Steinhauer, K., Mecklinger, A., Friederici, A. D., & Meyer, M. (1997). Probability and strategy: An event-related potential study of processing syntactic anomalies. *Zeitschrift fuer Experimentelle Psychologie, 2*, 305-331.

Steinhauer, K., Nadeau-Noel, K., Drury, J. E., & Royle, P. (2008). The priming of priming: A cross-linguistic ERP study. *Journal of Cognitive Neuroscience, G51*, 216.

Tockowitz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event related potential investigation. Studies in Second. *Language Acquisition, 2*, 173-204.

Trudeau, N., Frank, H., & Poulin-Dubois, D. (1999). Une adaptation en français du MacArthur Communicative Development Inventory. *Revue d'orthophonie et d'audiologie, 23*(2), 61-73.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270.

van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in Language Perception: Mild and Strong Conflicts Elicit Different ERP Patterns. *Journal of Cognitive Neuroscience, 22*(1), 67-82.

van den Brink, D., & Hagoort, P. (2004). The Influence of Semantic and Syntactic Context Constraints on Lexical Selection and Integration in Spoken-Word Comprehension as Revealed by ERPs. *Journal of Cognitive Neuroscience, 16*(6), 1068.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open and closed class words. *Journal of Memory and Cognition, 19*, 95-112.

Wassenaar, M., & Hagoort, P. (2007). Thematic role assignment in patients with Broca's aphasia: Sentence–picture matching electrified. *Neuropsychologia, 45*, 716-740.

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and Hearing Meaning: ERP and fMRI Evidence of Word versus Picture Integration into a Sentence Context. *Journal* *of Cognitive Neuroscience, 20*(7), 1235-1249.

Table 1

Expected components and task effects to different incongruent/discordant experimental conditions relative to the control sentence

Condition	Component(s)	Expected Task Effect
Auditory visual semantic mismatch	N400	None
Je vois un chapeau vert sur la table		
Auditory: I see a green hat on the tabe		
Visual: [green fish on table]		
Gender mismatch on adjective	LAN ?	None
*Je vois un chapeau *verte sur la table*	N400 ?	None
Auditory: I see a hat green.f on the tabe	P600	Stronger with GJ Task
Visual: [green hat on table]		

Table 2

Stimuli properties of Nouns and Adjectives used in the task: averages and standard deviations (in parentheses).

	Adjectives			Nouns		
	Variable	Invariable	T-test	Feminine	Masculine	T-test
AOE	26.33 (1.53)	23.75 (2.50)	n.s.	22.83 (2.48)	20.50 (2.95)	n.s.
Phon.	3 (0)	3.25 (0.50)	n.s.	4.16 (1.33)	4.50 (0.84)	n.s.
Syll.	1 (0)	1 (0)	_	1.50 (0.55)	1.83 (0.41)	n.s.
LexFreq	27.35 (25.37)	49.66 (31.19)	0.08	23.76 (24.50)	60.20 (93.04)	n.s.
LemmFreq	84.86 (52.25)	61.50 (53.35)	0.06	35.07 (42.80)	69.48 (96.02)	n.s.

Notes: AOE = age of emergence (50% of children) in months (Trudeau et al., 1999); Phon = length in phonemes; Syll. = length in syllables (Quebec French syllabification); LexFreq = Oral lexical frequency (films); LemmFreq = Oral lemma frequency (films) (both from New et al., 2001)

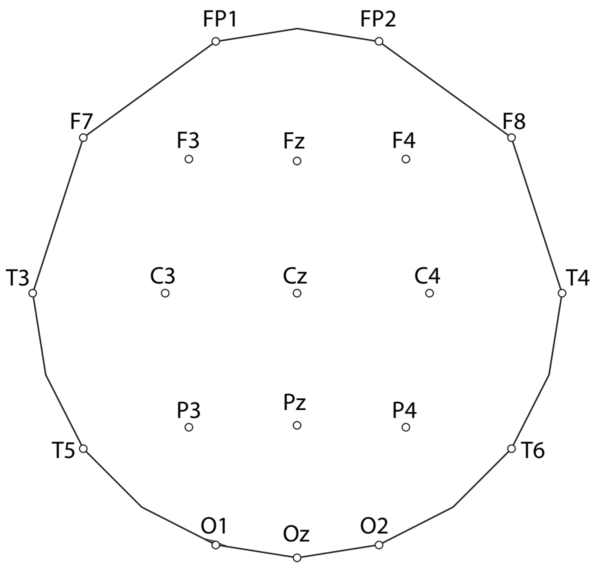
Table 3: Predicted and observed components and task effects to different experimental conditions relative to the control sentence.

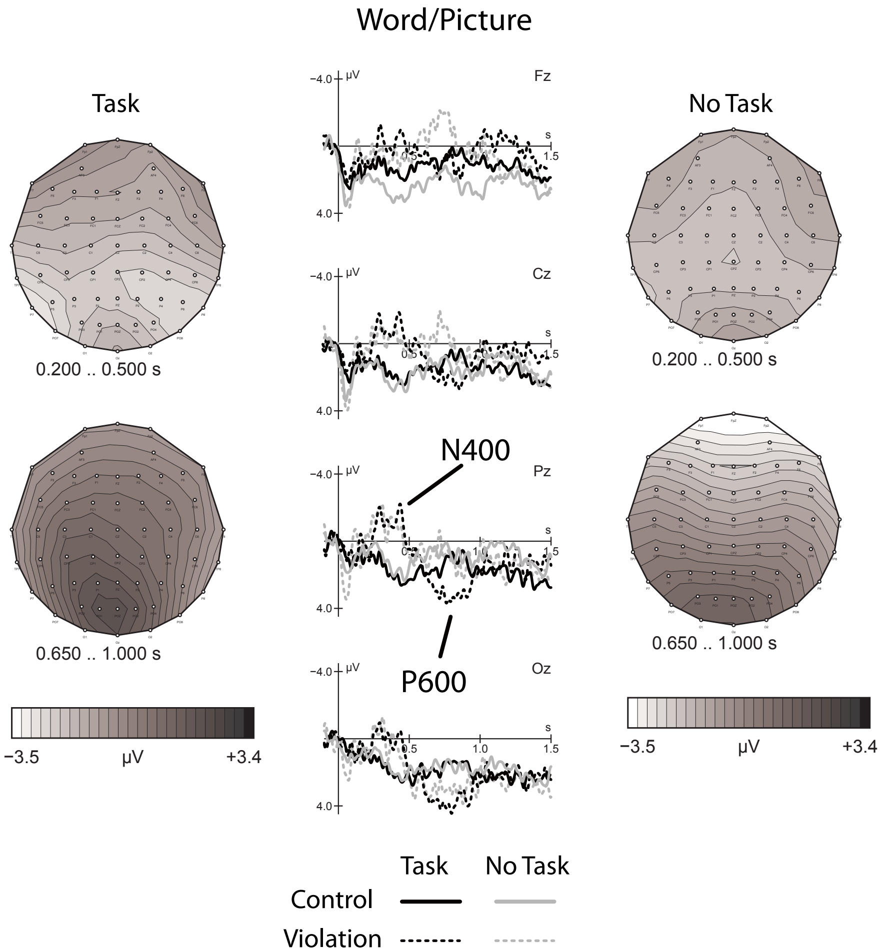
		Task effects	
Condition	Component(s)	Expected	Observed
Semantic mismatch	N400	None	None
	P600		Yes
Gender mismatch on adjective	LAN	None	Yes, focalization in No-task
	P600	Stronger in Task	Yes

Figure 1: Electrode map showing electrode array used for the experiment.

Figure 2: ERP waves at midline electrodes and voltage maps for the Word/Picture (semantic) mismatch conditions for both groups, time-locked to the critical word. Solid lines represent control conditions and dotted lines the mismatch ones. ERPs for the Task group are depicted in black, those of the No -Task group are in grey. Negative polarity is plotted upwards. Voltage maps present negativities in light grey and positivities in black. The top maps illustrate N400 effects, the bottom ones the P600 effects.

Figure 3: ERP waves at midline electrodes and voltage maps for the Noun-Adjective Gender mismatch conditions in both groups. Solid lines represent the control condition and dotted lines the gender agreement error condition. ERPs for the Task group are depicted in black, those of the No -Task group are in grey. Negative polarity is plotted upwards. Top maps illustrate the LAN effects, bottom maps illustrate P600s.


FP1
FP2
F7
F3
Fz
F4
F8
T3
C3
Cz
C4
T4
P3
Pz
P4
T5
T6
O1
Oz
O2



Word/Picture
Task
No Task
0.200 .. 0.500 s
0.650 .. 1.000 s
−3.5
µV
+3.4
Fz
Cz
Pz
Oz
−4.0
4.0
s
0.5
1.0
1.5
N400
P600
Control
Violation

Noun-Adjective Gender