Between alpha and gamma oscillations: Neural signatures of linguistic predictions and listener's attention to speaker's communication intention

Charlotte Stinkeste, Marion Vincent, Laurence Delrue, Angele Brunelliere

Title: Between alpha and gamma oscillations: neural signatures of linguistic predictions and listener's attention to speaker's communication intention

Authors: Charlotte Stinkeste[1], Marion A. Vincent[1], Laurence Delrue[2], Angèle Brunellière[1]

[1] Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France

[2] Univ. Lille, CNRS, UMR 8163 - STL - Savoirs Textes Langage, F-59000 Lille, France

Please address correspondence to:
Angèle Brunellière
SCALab, CNRS UMR 9193, Université de Lille, Domaine Universitaire du Pont de Bois, BP 60149, 59653 Villeneuve d'Ascq, France
Tel: (+33)3 20 41 72 04
angele.brunelliere@univ-lille.fr

Abstract: When listeners hear a message produced by their interlocutor, they can predict upcoming words thanks to the sentential context and their attention can be focused on the speaker's communication intention. In two electroencephalographical (EEG) studies, we investigated the oscillatory correlates of prediction in spoken-language comprehension and how they are modulated by the listener’s attention. Sentential contexts which were strongly predictive of a particular word were ended by a possessive adjective either matching the gender of the predicted word or not. Alpha, beta and gamma oscillations were studied as they were considered to play a crucial role in the predictive process. While evidence of word prediction was related to alpha fluctuations when listeners focused their attention on sentence meaning, changes in high-gamma oscillations were triggered by word prediction when listeners focused their attention on the speaker's communication intention. Independently of the endogenous attention to a level of linguistic information, the oscillatory correlates of word predictions in language comprehension were sensitive to the prosodic emphasis produced by the speaker at a late stage. These findings thus bear major implications for understanding the neural mechanisms that support predictive processing in spoken-language comprehension.

Count: 189 words (max.: 250 words)
Keywords: EEG, oscillations, prediction, communication intention, attention, sentence processing

1. Introduction

In spoken-language comprehension, listeners decode words from incoming speech, then quickly and incrementally elaborate sentence meaning thanks to semantic and syntactic unification (Hagoort, 2003). Successful comprehension rests on the ability to update sentence meaning after the recognition of each word. This mechanism is assumed to be optimized by top-down predictions. Listeners have been proposed to predict within different levels (e.g. at a semantic level through spreading activation) or between different levels of representations, from higher levels to lower levels such as the lexical level, during language comprehension (Pickering, & Gambi, 2018, for a review). We investigated prediction from sentence level to lexical level, as such a hierarchy between different levels of representations is also in line with the predictive coding framework (Friston, & Kiebel, 2009).

Based on the temporary representations of sentence meaning driven by prior context, lexical top-down predictions refer to the pre-activation of upcoming words before they are heard (see Kuperberg, & Jaeger, 2016; Pickering, & Gambi, 2018, for reviews). According to the predictive coding framework (Friston, & Kiebel, 2009), the brain continuously infers the probabilities of sensory input across the hierarchy of multi-level representations in order to predict upcoming input. Prediction error signals occur when the predicted information does not match the incoming information; they are used to update representations to ensure that future predictions are relevant. Like the predictive coding framework, lexical top-down predictions are considered to depend on the statistics of the linguistic environment (Kuperberg, & Jaeger, 2016; Levy, 2008). Predictive processing in sentences is thus explored by using a cloze probability test to measure the expectancy of the target word within a sentence (e.g. DeLong et al., 2005; Federmeier et al., 2007; Fleur et al., 2020; Foucart et al., 2014; Foucart et al., 2015; Li et al., 2017; Otten et al., 2007; Rommers, & Federmeier, 2018; Van Berkum et al., 2005; Wang et al., 2018; Wicha et al., 2003; Wicha et al., 2004). During a cloze probability test, participants are asked to complete a sentence frame with the first word that comes to mind: the expectancy of a word is defined as the proportion of participants who choose that same word to complete the sentence. Although evidence of predictive processing during the presentation of both written sentences (DeLong et al., 2005; Fleur et al., 2020; Foucart et al., 2014; Wicha et al., 2004) and spoken sentences (Foucart et al., 2015; Otten et al., 2007; Van Berkum et al., 2005; Wicha et al., 2003) has been widely explored by recording electrophysiological brain activity, oscillatory activity related to linguistic predictions and its interaction with listener's attention are rarely considered.

Since predictive processing in language comprehension requires strong efficient inter-network transfer of information across distributed brain regions which are specialized in the processing of various levels of representations (sentence level vs. word level), oscillatory correlates appear to be a plausible neural signature of linguistic predictions in interaction with listener's attention. Attention is known to play a crucial role in language comprehension, as shown by a recent functional magnetic resonance imaging (fMRI) study revealing that sentence-level syntactic and semantic unification is prevented in situations of inattention (Cohen et al. 2021). Similarly, Boudewyn, & Carter (2018) highlighted a mapping between alpha-band oscillations (8–12 Hz), which are associated with attentional engagement, and successful listening comprehension. In non-linguistic domains, predicting upcoming information shifts the attention to this expected event, leading to the suppression of alpha activity related to a strong attentional engagement (Foxe et al., 1998; Mayer et al., 2016).

Several electrophysiological studies have focused on linguistic predictive processing in sentences while quantifying oscillatory correlates (e.g. Brunellière et al., 2022; Li et al., 2017; Molinaro et al. 2017; Piai et al., 2014, 2015; Rommers et al., 2017; Wang et al., 2018). Two recent review papers (Meyer, 2018; Prystauka, & Lewis, 2019) revealed the strong involvement of beta and gamma oscillations in top-down predictions during sentence processing in addition to alpha activity. While beta oscillations (16-20 Hz) seem to be associated with maintaining information in the working memory and reflect the top–down propagation of predictions to lower processing levels, gamma oscillations (around 30 Hz or between 60 and 100 Hz) are more related to the matching between incoming input and top-down predictions (Lewis, & Bastiaansen, 2015). Reduced beta power was observed during the processing of sentential contexts which were highly predictive of an upcoming word in comparison with weakly predictive sentential contexts (Li et al., 2017; Piai et al., 2014, 2015; Rommers et al., 2017; Wang et al., 2018). Low gamma activity around 30 Hz (i.e. lower than 45 Hz, see Penolazzi et al., 2009; Weiss, & Mueller, 2003) reflects the matching between incoming input and top-down predictions; gamma-band power is thus higher with semantically congruent sentences than with semantically incongruent ones (Hald et al., 2006; Penolazzi et al., 2009; Rommers et al., 2013; Schneider et al., 2008). Although high-gamma activity is also related to the matching between incoming input and top-down predictions, it has been found to be higher with semantically incongruent words than with semantically congruent ones (Wang et al., 2018). Some authors have already interpreted high-gamma activity as reflecting the propagation of prediction errors in language comprehension and in speech perception (Arnal, & Giraud, 2012; Lewis, & Bastiaansen, 2015).

Few electrophysiological studies have investigated oscillatory correlates of prediction in language comprehension after the processing of adjectives preceding predicted nouns (Brunellière et al., 2022; Molinaro et al. 2017). Although oscillatory correlates of prediction in language comprehension have been studied by looking at the processing of expected and unexpected nouns embedded in a sentential context or by comparing the processing of highly semantically constraining sentences to that of weakly semantically constraining sentences, the most reliable method to provide a conclusive interpretation of lexical top-down predictions is probing the brain's reaction to adjectives or determiners either matching the linguistic properties of the predicted words they precede or not. As initially suggested by Kutas & Hillyard (1980), incoming words may be easier to integrate into the representation of sentence meaning depending on how well their meaning fits with it. Therefore, brain responses to incoming words embedded in a prior sentential context may be the result of top-down predictions for upcoming words or the integration of the incoming words into sentence representation. Indeed, evidence of prediction effects in sentence processing requires identifying them before the occurrence of the expected words (Pickering, & Gambi, 2018), in order to avoid inconclusive interpretations due to the integration of the incoming words into sentence representation. To provide clear evidence of word prediction effects, Delong et al. (2005) and Wicha et al. (2003, 2004) designed experiments with an adjective or a determiner preceding the word predicted from prior sentential context and having gender or phonological properties either matching those of the predicted word or not. Importantly, the meanings of the adjective or the determiner preceding the predicted word did not differ. One may thus exclude the idea that the processing of the words preceding the predicted one is more difficult to integrate into a given sentence representation. Using the same design, Molinaro et al. (2017) showed lower beta power for gender mismatching adjectives than for those matching the predicted nouns, while these adjectives were visually presented on the screen in a word-by-word presentation. In contrast, Brunellière et al. (2022) found a prediction effect over the low-gamma band with auditorily presented adjectives after the audiovisual presentation of sentential context (face of the speaker accompanied by their voice). The differences in neural signatures of oscillatory correlates may be explained by differences in the speed of exposure to sensory information: a written word-by-word presentation is a slow exposure of words within a sentence compared to a natural exposure of audiovisual sentences. Beyond this potential explanation, the two electrophysiological studies investigating oscillatory correlates of prediction effect in language comprehension after the processing of adjectives preceding predicted nouns (Brunellière et al., 2022; Molinaro et al. 2017) did not report the same correlates. The neural signatures of prediction effect in sentence processing thus remain elusive to this date.

To better understand these neural signatures of linguistic prediction in sentence processing, we investigated the oscillatory correlates of predictions in spoken language comprehension and how they are modulated by listener's attention. Similar to Brunellière et al. (2022) and Molinaro et al. (2017), we focused on the pre-activation of lexical representations prior to that predicted word appearing in the input as being evidence for prediction in sentence comprehension. As highlighted by Pickering & Gambi, (2018), clear evidence for word prediction in sentence comprehension requires identifying it before the occurrence of the predicted word.

According to the attention allocation account, listeners pay attention to the parts of the utterance where the accent falls (Cutler, 1976; Sanford, Sanford, Molle, & Emmott, 2006). The accented speech thus acts as a salient external stimulus capturing exogenous attention. Since attention is allocated to accented words, better retention in memory was found for accented words embedded in sentential context (Birch, & Garnsey, 1995; Blutner, & Sommer, 1988; Fraundorf, Watson, & Benjamin, 2010). Prosodic emphasis of the sentential context also triggered deeper lexical-semantic processing of the following words. The amplitude of the N400, which is the electrophysiological component associated with lexical-semantic processing, was increased after semantically incongruent words when they were embedded in a sentential context produced with natural prosodic emphasis (Brunellière et al., 2019). Some authors have thus argued that allocating attention to accented words produces more detailed semantic processing (Birch, & Garnsey, 1995; Blutner, & Sommer, 1988; Fraundorf, Watson, & Benjamin, 2010; Li, & Lu, 2011). Despite the beneficial effect of accented speech capturing listener's attention on the processes of language comprehension, the impact of listener's attention via prosodic emphasis on oscillatory correlates of predictions in spoken language comprehension is yet to be investigated.

In the present study, we explored listener's attention in two different ways by focusing on exogenous and endogenous attention. This distinction was defined by Posner & Cohen in 1984. Exogenous attention refers to attention driven by external events, while endogenous attention is driven by the goal of tasks. In two electroencephalographical (EEG) experiments, we questioned the hypothesis of exogenous attention captured by prosodic emphasis of the sentential context. The speaker either highlighted or not the content of the message to convince the interlocutor, leading to natural prosodic emphasis of the sentential context or not. This prosodic manipulation was used to investigate how the natural prosodic emphasis is integrated with linguistic prediction in sentence processing and whether this interaction is related to the hypothesis of exogenous attention to provide a better understanding of the links between exogenous attention and prediction. The impact of endogenous attention was investigated by changing the goal of the tasks in each experiment. In Experiment 1, participants were asked to listen to spoken sentences and understand their meaning. In Experiment 2, they were asked to judge the intention of the speaker. In the latter experiment, endogenous attention was on the pragmatic information provided by prosodic cues. In both experiments, we presented semantically constraining spoken sentences followed by a possessive adjective that either matched the gender of the expected (albeit not presented) word or not (see Figure 1), in order to probe word prediction effects.

< Insert Figure 1 here >

Regarding the neural signatures of predictions in spoken-language comprehension, we explored oscillatory activity in the alpha, beta and gamma frequency bands during the processing of critical adjectives preceding predicted words, providing direct evidence of word prediction based on prior sentential context. In two experiments, we investigated the oscillatory correlates of predictions in spoken language comprehension and how they are modulated by the listener's exogenous attention via prosodic emphasis of the sentential context. Alpha activity is assumed to play a role between attention and prediction while prediction appeared to elicit an attentional shift to the predicted information (Foxe et al., 1998; Mayer et al., 2016). In line with these studies (Foxe et al., 1998; Mayer et al., 2016), we expected a decrease in alpha activity after gender-matching adjectives in comparison to gender-mismatching adjectives owing to higher attentional engagement towards the former. According to the attention allocation account claiming that exogenous attention captured by prosodic emphasis causes beneficial effects on the processes of language comprehension, we hypothesized that the prosodic emphasis of sentential context would trigger stronger expectations on the predicted word, leading to more exogenous attentional engagement towards gender-matching than towards gender-mismatching adjectives. As the alpha activity indexes attentional engagement towards the expected event (Foxe et al., 1998; Mayer et al., 2016, Strube et al., 2021), we expected a greater decrease in alpha activity after gender-matching adjectives than after gender-mismatching adjectives when the context was produced with emphasis. In line with previous studies (Molinaro et al., 2017; Brunellière et al., 2022), an alternative hypothesis is that beta and gamma oscillations may also be the oscillatory correlates of predictions in language comprehension. Since studies by Brunellière et al. (2022) and Molinaro et al. (2017) showed that predictions occur only in more challenging situations, beta and gamma oscillations related to the prediction effect may be observed only when there is no prosodic emphasis of sentential context. Previous research has highlighted that when the speech signal is degraded, listeners rely more on contextual information provided by the sentence for successful comprehension (Bhandari, Demberg, & Kray, 2022; Obleser, & Kotz, 2010; Obleser et al., 2010; Sheldon et al., 2008). Although the lack of prosodic emphasis is not comparable to when the speech signal is degraded, the incoming signal may be less reliable owing to it, leading to more use of the sentential context to support language comprehension.

Moreover, we examined whether the neural signatures of predictions in spoken-language comprehension differed as a function of endogenous attention on sentence meaning (Experiment 1) or on pragmatic information (Experiment 2). For example, if changes in alpha oscillations are associated with the specific links between exogenous attention and predictions, these changes should be observed independently of endogenous attention to a particular linguistic level. In both experiments, we thus expected a greater decrease in alpha activity associated with a main effect of prediction and interactive effects between exogenous attention and prediction.

2. Methods: Experiment 1

2.1 Participants

Thirty-two native[1] French students from the University of Lille participated in Experiment 1 (mean age = 21 years old, range: 18–28 years old, standard deviation: 2.6 years old; 3 males). None of them had hearing, visual, language or neurological impairments. They were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and they self-reported that they were not taking any medication. They received monetary compensation for their participation and gave their written informed consent before the beginning of the experiment. Six additional participants were excluded during the EEG pre-processing stage owing to excessive blinking and movement artifacts. The study (including Experiments 1 and 2) was approved by the Research Ethics Committee of the University of Lille.

[1] All participants had French as their native language and started learning other languages only in primary school.

2.2 Stimulus

The stimuli consisted of a set of 160 pairs of strongly semantically constraining sentence frames (mean cloze probability: 0.73; range: 0.6-1; standard deviation: 0.15). For example, the following sentence frame "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi…*" (in English, "You helped me when I was undecided. I followed...") was highly predictive of the final noun "*conseil*" (i.e. in English, "advice") (see, Figure 1). The semantic constraints of the sentence frames were assessed by a classical cloze probability test in which French speakers (who did not take part in Experiments 1 and 2) were asked to complete sentential contexts with the first word that came to their mind and match it with an adjective when that word was a noun. Each sentence frame was given to 15 participants. As in previous studies (Brunellière, & Soto-Faraco, 2013, 2015; Brunellière, Vincent, & Delrue, 2022; Otten et al., 2007; Van den Brink, Brown, & Hagoort, 2001; Van den Brink, & Hagoort, 2004), a sentence frame was defined as highly predictive if at least 50% of the participants gave the same word to complete it.

The mean length of strongly semantically constraining sentence frames was 17.7 words (range: 9-18 words, standard deviation: 3.8 words) and they were always followed by a possessive adjective ("*ton*", the masculine form, or "*ta*", the feminine form, your in English), which involved the listener in the listening process. This adjective expresses the second person singular, so that the listener may infer that this message was dedicated to them. Thanks to this adjective, we manipulated gender congruency between the predicted noun and the adjective (i.e. either unexpected gender, UG, or expected gender, EG, with respect to that of the predicted noun, see Figure 1). The gender (feminine or masculine) of the adjective either did not match (unexpected gender, UG) or matched (expected gender, EG) that of the predicted noun. For instance, the following sentential context: "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi…*" predicted the masculine word "*conseil*" (see Figure 1). If the masculine form of the possessive adjective referring to the second person, "*ton*", was the incoming word after the following sentential context: "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi*…", the gender of the adjective would match (expected gender, EG) that of the predicted noun. In contrast, if the feminine form of the possessive adjective referring to the second person, "*ta*", was the incoming word after the same sentential context, the gender of the adjective would not match (unexpected gender, UG) that of the predicted noun. The classical cloze procedure made it possible to check that the use of the possessive adjective was adequate and plausible in our sentences. All adjectives began with an occlusive consonant with a prominent onset (/t/), in order to precisely determine the onset of adjectives in the auditory recordings. The nouns predicted by the strongly semantically constraining sentence frames were singular nouns with two or three syllables selected from the Lexique database (New, Pallier, Brysbaert, & Ferrand, 2004). The 160 predicted nouns were equally divided into 80 feminine and 80 masculine words. All predicted nouns were familiar words (mean frequency per million occurrences: in film subtitles = 16.05; in books = 14.75, range per million occurrences: 1-392; 1-233, standard deviation per million occurrences: 43.5; 28.8).

In addition to gender congruency, the second variable we manipulated was the prosodic emphasis of sentential context: the sentence was either produced with emphasis (E+, i.e. the speaker wants to convince the listener) or without emphasis (E-, i.e. there was no intention to convince the listener). There were thus four experimental conditions in the experimental design (see Figure 1) with a total of 160 strongly semantically constraining sentence frames (40 trials per condition). In order to prevent strategies based on the structure of the critical sentences, 80 filler sentences were also used. There were as many weakly semantically constraining filler sentences as there were critical sentences from which prediction errors could be triggered. All filler sentences were weakly semantically constraining sentences and the possessive adjective ("*ta*" or "*ton*") was either placed at the beginning or in the middle of these sentences. This enabled us to avoid strategies on the possessive adjective, which was always placed in the last part of the critical sentences.

Regarding the recording of the stimuli, a male French speaker (a student from the University of Lille unknown to the participants and aged 36 years old) in a deaf room facing a unidirectional microphone was asked to produce each sentence with natural prosody at a normal speaking rate. A teleprompter was placed in front of him and he had to pronounce each sentence twice. All sentences sharing the same sentence frame were successively spoken and the order of the four versions within one sentence frame varied for each of them to prevent the effect of first reading in a particular condition. As we were studying the ecological phenomenon of prosodic emphasis, we did not ask the speaker to stress a specific point in the sentence. He was asked to pronounce the sentences naturally, either with emphasis or with no specific emphasis as in neutral prosody. In the present study, the expressive function was thus used and pragmatic information could be extracted from prosodic cues.

< Insert Figure 2 here >

Since strongly semantically constraining sentence frames were composed of a grammatical cataphora, the opening of a parenthesis marked by a prosodic boundary was naturally produced before the production of the final word with its adjective (see Figure 2). A grammatical cataphora is the use of a pronoun to refer to another word ahead of it in a sentence (called, a postcedent). For example, in the sentence frame "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi*……", meaning "You helped me when I was undecided. I followed…", the pronoun « *l'* » refers to the predicted word "*conseil*" ("advice", in English). This grammatical cataphora places the predicted word with its possessive adjective in the position of a parenthetical element. As seen in Figure 2, this creates a pause before the predicted word with its possessive adjective, thus isolating it from acoustic variations in the prosodic emphasis of speakers during the production of sentential context. This isolation of the predicted word and its possessive adjective from the sentential context was evidenced by the fact that no significant difference in fundamental frequency, intensity and duration[2] was found during the production of the adjective as a function of the Congruency of gender and Prosodic emphasis factors (mean fundamental frequency, EGE-: 104 Hz, UGE-: 103 Hz, EGE+: 105 Hz, UGE+: 106 Hz, no main effects of Congruency of gender and Prosodic emphasis, Congruency of gender, $F(1,159)=0.09$, $p>0.2$, Prosodic emphasis, $F(1,159)=1.63$, $p>0.2$, no interaction between Congruency of gender and Prosodic emphasis, $F(1,159)=0.17$, $p>0.2$; mean intensity, EGE-: 49 dB, UGE-: 49 dB, EGE+: 49 dB, UGE+: 49 dB, no main effects of Congruency of gender and Prosodic emphasis, Congruency of gender, $F(1,159)=0.01$, $p>0.2$, Prosodic emphasis, $F(1,159)=1.68$, $p>0.2$, no interaction between Congruency of gender and Prosodic emphasis, $F(1,159)=0.07$, $p>0.2$, duration, EGE-: 0.074 s, UGE-: 0.074 s, EGE+: 0.077 s, UGE+: 0.075 s, no main effects of Congruency of gender and Prosodic emphasis, Congruency of gender, $F(1,159)=0.48$, $p>0.2$, Prosodic emphasis, $F(1,159)=1.94$, $p>0.2$, no interaction between Congruency of gender and Prosodic emphasis $F(1,159)=0.52$, $p>0.2$). Importantly, by asking the speaker to pronounce the sentences naturally, either with a strong intention or with no intention to convince, words were more emphasized in the sentential context when the speaker wanted to convince the listener. As shown in Figure 2, the highest peaks of fundamental frequency and intensity occurred in the same words. In line with the aim of the study, the sentence context spoken with intended emphasis was marked by an increase in fundamental frequency and intensity in comparison with the sentence context spoken without prosodic emphasis. This was evidenced by significant differences in fundamental frequency and intensity during the production of sentential contexts (mean fundamental frequency, E+: 155 Hz, standard deviation, SD: 11 Hz, range: 126-186 Hz, E-: 124 Hz, SD: 7 Hz, range: 106-143 Hz, $t(318)=29.25$, $p=6.9\times10^{-7}$; mean intensity, E+: 64 dB, SD: 1.9 dB range: 59-69 dB, E-: 56 dB, SD: 1.9 dB, range: 51-61 dB, $t(318)=36.89$, $p=1.2\times10^{-7}$). Unlike these two acoustic correlates and what can be seen in Figure 2, there was

[2] All acoustical measurements were extracted using the Praat speech editing software (version 5.3; Boersma, &Weenink, 2011)

no significant difference in duration between sentential contexts spoken with prosodic emphasis or without it (E+: 3.393 s, SD: 863 ms, range: 1497-6501 ms, E-: 3.276 s, SD: 855 ms, range: 1551-6377 ms, $t(318)=1.21$, $p>.2$). As described in prior literature about French prosody (Jun, & Fougeron, 2002; Lacheret-Dujour, & Beaugendre, 1999; Pasdeloup, 1990; Touati, 1987), the sentential contexts spoken with intended emphasis showed increased fundamental frequency and intensity in our experiments.

The recordings were digitized at a sampling rate of 44,100 Hz at 16 bits. To ensure that intonation, speaking rate and duration were equivalent up to the word before the target adjective for a sentence frame produced with a particular emphasis, we used a splicing technique. A recording of each sentence frame up to the word just before the target adjectives was selected by two native speakers (e.g. in the example shown in Figure 1B, "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi*…", You helped me when I was undecided. I followed…), so that the intonation and speaking rate sounded natural and all words were pronounced clearly. The splicing technique was performed separately at each degree of prosodic emphasis (with emphasis vs. without it). Sentential contexts were acoustically identical within one degree of prosodic emphasis and the ecological phenomenon of prosodic emphasis was preserved. Fragments coming from other recordings of the same sentence frame (e.g. in the example, "*ta conseil*" your advice, "*ton conseil*" your advice) then completed the sentence to create two new versions: one containing the expected adjective and the other containing the unexpected adjective. As in Brunellière et al. (2022) and Foucart et al. (2015), the predicted final nouns were never presented after the adjectives. This avoided any interference effect between the processing of adjectives and that of predicted words, owing to strategies related to linguistic violations. Using Cool Edit, the predicted final nouns were replaced by a brown noise at the mean intensity of the two adjectives ("*ta*" or "*ton*") for each sentence frame. The duration of auditory distortions was identical for one sentence frame with a mean duration of 1 s (range:

0.7-1.54 s). The mean duration of adjectives was 0.075 s (range: 0.03-0.136 s, standard deviation: 0.015 s) and did not vary significantly between the expected and unexpected gender conditions (p>.2). Filler sentences had the same distortions as critical sentences, although the duration of distortions varied between five different values (0.04, 0.1, 0.2, 0.3 or 0.4 s) and were never placed in the final part of the sentences. In the filler sentences, the distortions came at random places before the final part and were never placed just after the adjectives ("*ta*" or "*ton*").

2.3 Experimental procedure

In Experiment 1, participants were asked to listen to sentences for comprehension (for similar approaches, see Brunellière et al., 2019, 2020, 2022; Van Den Brink, & Hagoort, 2004) and they were informed that they had to perform a lexical recognition task after listening to sentences (see Figure 1). From the beginning of the experiment, participants were informed that after the listening phase, they had to perform a lexical recognition task in which they had to indicate whether they had been previously exposed to these words. As in previous studies focusing on linguistic predictions in spoken-language comprehension (e.g. Brunellière et al., 2022; Foucart et al. 2015), this task ensured that the meaning of sentential contexts had been well computed by observing more false alarms for the predicted words than for new words that were not exposed or expected while listening to sentence frames. Participants were not instructed about the manipulation of gender congruency or prosodic emphasis. However, they were informed that the sound signal could be interrupted and noisy in the sentence listening task, as in everyday life telecommunication problems.

2.3.1 Sentence listening task

Participants were placed in front of a screen and a response box in a Faraday cage. Their electrical brain activity was recorded during the sentence listening task. As illustrated in Figure 1A, a red fixation cross in the center of the screen was first presented on a black background for 0.5 s before the onset of the auditory sentence and remained until after the end of it. To minimize artifacts, participants were asked to focus their gaze on the center of the screen and to keep their eyes as still as possible. They were encouraged to avoid moving unless the grey fixation cross was displayed. A black screen then appeared for 1 s and was replaced by a grey fixation cross at the center of the monitor for 2 s before the presentation of the next trial. The auditory stimuli were played binaurally at a comfortable sound level via headphones. Participants listened to 16 practice sentences prior to the set of four 10-min blocks of 60 trials, each containing sentences from all experimental conditions and fillers presented in random order.

2.3.2 Lexical recognition task

During both listening and lexical recognition tasks, participants sat in a shielded room. Participants were asked to indicate, as quickly and accurately as possible, whether they had heard the word or not during the listening task by pressing one of the two buttons on a response box. These response buttons were counterbalanced across participants. Each word among a set of 320 words was presented randomly at the center of the screen. For each participant, half of the words never appeared in the sentences during the listening task (160), and the remaining half came from each experimental condition and from fillers (160). Among the words which never appeared during the listening task, half of them were new (80) and the other half was expected from the sentence frame, yet muted (80). Among the latter, 20 words were expected from each experimental condition and they were equivalent in psycholinguistic properties (lexical frequency, length in letters and phonemes and phonological and orthographical neighborhood, see supplementary material, Table 1). Depending on the experimental list of the listening task, each word was attributed to one experimental condition (EGE-, UGE-, EGE+, UGE+). A fixation cross appeared on the screen 0.5 s before the presentation of each word and a black screen was present for 1 s after the participants’ responses and before the next trial (see

Figure 1C).

2.4 Data recording, pre-processing and analyses

2.4.1 Behavioral data

The participants' behavioral data were recorded with the E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) using a response box in the lexical recognition task. According to the signal detection theory (DeCarlo, 1998), we computed individual values of *d*-prime using hit responses (words heard during sentence listening and participant pressing the button corresponding to 'heard words') and false alarms (foils never presented during the sentence listening phase, but for which participants pressed the button corresponding to 'heard words'). A Student's *t*-test was performed on individual values of *d*-prime to check whether participants performed lexical recognition task above chance level. A repeated measures ANOVA was also performed comparing false alarms to examine whether there were more false alarms for expected yet unheard words than for new words. This analysis included one factor with five modalities (EGE-, UGE-, EGE+, UGE+, new words).

2.4.2 EEG data

The electrical signal from the scalp (sampling rate: 1024 Hz) was recorded during the sentence listening task with a 128-channel BioSemi Active Two AD-box system. Two external electrodes measured eye movements from the right eye (vertical and horizontal electro-oculograms) and two other external electrodes were used to measure the activity over the right and left mastoids (later used as off-line reference). As recommended with the Biosemi Active Two AD-box, individual electrodes were adjusted to a stable offset lower than 20 mV during the EEG recording. Artifact rejection was performed using the Cartool software (https://sites.google.com/site/cartoolcommunity/home) under a rejection criterion of 70 µV for any channel, in a segment starting 2 s before the onset of the adjectives and after the onset of sentences and ending 2 s after them. Segments with an amplitude difference between two consecutive EEG time frames which was superior to 70 µV over one electrode from the scalp or external electrodes were rejected. Blinks and eye movements as well as other muscle artifacts were therefore removed. The number of accepted trials did not differ between all four experimental conditions with an average of 37.7 accepted trials ($p>.2$; EGE-, mean: 37.7, standard deviation, SD: 2.3; UGE-, mean: 37.8, SD: 1.8; EGE+, mean: 37.4, SD: 2.3; UGE+, mean: 37.7, SD: 2.1). Time-frequency and statistical analyses on remaining trials after the artefact rejection step were performed with Matlab, using the open-source toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The individual EEG channels were re-referenced to the average of the signals recorded at the two mastoids. Raw EEG data were then segmented and time-locked to the onset of sentences and adjectives (-2 s to + 2 s). The segments time-locked to the onset of sentences were used to employ a baseline which enabled to rule out any bias due to the processing of sentential context. This baseline occurred in a time window between -0.50 and -0.20 s relatively to sentence onset, as previously used by Brunellière et al. (2022) and Wang et al. (2018). A bandpass filter of 0.1 to 100 Hz (2nd-order Butterworth filter, 0.01 – 100 Hz) was also applied.

Following the procedure by Brunellière et al. (2022) and Wang et al. (2018) time-frequency representations (TFRs) were computed for each participant, channel and segment, in two overlapping frequency ranges to optimize the resolutions within the frequency and time domains. For low frequencies between 2 and 30 Hz, a 0.5-s Hanning window (single taper) was used to calculate the power changes with a frequency step of 2 Hz and a time step of 0.01 s. For high frequencies between 25 and 100 Hz, a Slepian multitaper was applied at a frequency step of 5 Hz, a time step of 0.05 s, 10-Hz frequency smoothing and a 0.2-s window. We divided the TFRs of adjective onset epochs in each experimental condition by the baseline, consisting in the average of all sentence onset epochs over a time window between -0.50 and -0.20 s relatively to sentence onset. A log transformation (10*log10) was then applied to power values to provide them in decibels. Similar to Brunellière et al. (2022) and Wang et al. (2018), we used a [-1; 1]-s time window time-locked to the target of interest, so that this window was long enough to observe brain activity after the processing of adjectives but shorter than that used for data segmentation to avoid a ringing artifact on the signal. Over this time window, cluster-based permutation statistics (Maris, & Oostenveld, 2007) were performed across participants over all 128 electrodes in four different frequency bands (alpha: 8 – 12 Hz, beta: 16 – 20 Hz, low gamma: 25 – 40 Hz, high gamma: 60 – 100 Hz). This non-parametric statistical procedure solves the multiple comparison issue optimally. To determine the oscillatory correlates of predictions in spoken language comprehension, we compared TFRs between the EG and UG conditions in the interval after the presentation of the adjective (0 to 1 s) over each frequency band. Based on our hypothesis, we also conducted a potential two-way interaction between the two factors (Congruency of gender and Prosodic emphasis) by performing permutation tests on the differences in TFRs for UG minus EG between the E- and E+ conditions. All statistical comparisons were quantified using a *t*-test and a threshold of 95^{th} quantile was applied to determine cluster candidates. Cluster-level statistics were computed by adding the *t*-values within each cluster. All adjacent data points according to the adjacent neighbors' design exceeding significance level (0.05 %) were grouped into clusters. Significance probability was calculated using the Monte Carlo method, with 1,000 random permutations.

3. Results: Experiment 1

3.1 Behavioral results

In the lexical recognition task, individual values of *d*-prime significantly differed from the null hypothesis of chance performance ($t(31)=5.69$, $p<.05$), showing that participants paid attention to sentence frames (mean value of *d*-prime: 0.30). The low value of *d*-prime was due to the high number of words expected from the sentence frames but never presented in the listening phase. An ANOVA analysis on false alarms including one factor with five modalities (EGE-, UGE-, EGE+, UGE+, new words) revealed a main effect of this factor ($F(4,124)=21.29$, $p=2.1\times10^{-13}$). By comparing the false alarms of one experimental condition to those of new words, post-hoc Tukey tests showed that participants produced more false alarms for expected yet unheard words than for new words (i.e. unexpected and unheard words, EGE-: 0.45, UGE-: 0.38, EGE+: 0.46, UGE+: 0.41, new words: 0.25, $ps=1.7\times10^{-6}$), suggesting that participants paid attention to the meaning of sentential contexts. There were no significant effects between experimental conditions on false alarms.

3.2 EEG results: Experiment 1

We compared the TFRs of alpha (8 – 12 Hz), beta (16 – 20 Hz), low gamma (25 – 40 Hz) and high gamma (60 – 100 Hz) band activities for the UG and EG conditions in the interval between 0 and 1s after adjective onset. Figure 3 shows a significant cluster for the alpha activity in the observed data between the UG and EG conditions over the frontocentral electrodes and in a time window between 0 and 0.23 s after adjective onset ($p=3.3\times10^{-2}$). A greater decrease in alpha power was elicited by the EG than by the UG condition (Figure 3B). Alpha activity was suppressed with respect to the baseline and this suppression was stronger in the EG condition than in the UG one.

< Insert Figure 3 here >

The cluster-based permutation tests on the differences in TFRs for UG minus EG between the E- and E+ conditions for the alpha activity revealed a significant cluster in the observed data over the central and right parietal electrodes and in a time window between 0.55 and 0.86 s after adjective onset (see Figure 4, $p=1.7\times10^{-2}$). In this cluster, alpha activity was higher in the UG than in the EG condition when the context was produced with emphasis ($p=1.7\times10^{-2}$). This was due to a greater decrease in alpha power in the EG than in the UG condition. No significant difference between the UG and EG conditions and no interactive effect between congruency of gender and prosodic emphasis were found over the beta, low gamma and high gamma frequency bands.

< Insert Figure 4 here >

4. Discussion: Experiment 1

In line with previous electrophysiological studies of sentence processing (e.g. DeLong et al., 2005; Fleur et al., 2020; Foucart et al., 2014; Foucart et al., 2015; Wicha et al., 2003; Wicha et al., 2004), the findings of Experiment 1 show clear signs of predictive processing in language comprehension. By examining the processing of critical adjectives preceding predicted words, we found evidence of word prediction based on prior sentential context. As previously described (see Kuperberg, & Jaeger, 2016; Pickering, & Gambi, 2018, for reviews), lexical top-down predictions refer to the pre-activation of upcoming words before they are heard and are based on the temporary representations of sentence meaning driven by prior sentential context. When upcoming words are pre-activated, their linguistic properties (phonological, semantic and grammatical) are also activated. Therefore, the difference in the brain response between gender-matching and gender-mismatching adjectives regarding a particular predicted word reflects the mismatch between the grammatical gender of the predicted word and that of the preceding adjective. This brain response may thus provide information on the elicitation of word prediction and indirectly relate to a sentence meaning component. Interestingly, changes in alpha activity were the oscillatory correlates of prediction in spoken language comprehension that we found in Experiment 1. It was not surprising to observe changes in alpha oscillations, which are known to be associated with attentional engagement and predictions (Foxe et al., 1998; Mayer et al., 2016). In line with our hypothesis based on previous studies in non-linguistic domains (Foxe et al., 1998; Mayer et al., 2016), we observed a greater decrease in alpha activity after gender-matching than after gender-mismatching adjectives. Based on these previous studies in non-linguistic domains (Foxe et al., 1998; Mayer et al., 2016) investigating alpha activity after expected and unexpected information, we interpreted a greater decrease in alpha activity after gender-matching adjectives as reflecting a higher attentional engagement towards them.

To our knowledge, this is the first time that changes in alpha activity have been observed when the processing of gender matching adjectives predicting words was compared to that of gender mismatching adjectives. While Molinaro et al. (2017) found lower beta power for gender mismatching than for gender matching adjectives in a word-by-word presentation, Brunellière et al. (2022) found a prediction effect over the low-gamma band with auditorily presented adjectives after the audiovisual presentation of sentential context. Beyond the differences in frequency bands for oscillatory correlates of prediction in language comprehension, Brunellière et al. (2022) and Molinaro et al. (2017) also highlighted evidence for predictive processing in challenging situations. In Experiment 1, the listeners did not have to deal with complex situations, including language comprehension in a second language or degraded incoming information. Instead, they were exposed to auditory sentences naturally produced by a native speaker, who expressed prosodic emphasis a third of the time. One may assume that this frequency is sufficient to guarantee strong motivation from the listeners to carry out a fine analysis of sentential context, leading to easy prediction of the upcoming words. Interestingly, no evidence of word prediction was found in both experiments conducted by Brunellière et al. (2022), in which the same linguistic sentence structures as in the present experiment were presented in auditory-modality alone, yet produced in neutral prosody. Therefore, the prosodic emphasis expressed in one third of the stimuli in Experiment 1 could be a sufficient frequency to guarantee strong motivation from the listeners to generate word predictions.

Moreover, we investigated whether lexical top-down predictions based on sentence meaning immediately integrate the aspects of prosodic emphasis and whether this is related to the hypothesis of exogenous attention, to throw light on the links between exogenous attention and prediction. An interesting finding of Experiment 1 relates to the time course of interactive effects between gender congruency and prosodic emphasis. Remarkably, these effects occurred later than the prediction effect alone, although the alpha activity indexing attentional engagement towards the expected event was the pattern found when the prediction effect was observed independently of prosodic emphasis over an early time window or when it interacted with it later. While a greater decrease in alpha power elicited by the expected gender than by the unexpected gender occurred in a time window between 0 and 0.23 s after adjective onset, an interactive effect between gender congruency and prosodic emphasis was evidenced later in a time window between 0.55 and 0.86 s after adjective onset. This suggests that prosodic emphasis was taken into account in a second step after an initial step where the prediction of upcoming words occurs on the bases of the semantic content of sentential context. The present results are relevant regarding the question of the moment when pragmatic computations, including the intentional aspect of communication and perspective-taking, are taken into account in spoken language comprehension. Two opposite pragmatic accounts of language comprehension have been proposed until now (Clark, 1996; Van Berkum et al., 2008). The one-stage account assumes that the meanings of individual words and pragmatic computations are immediately combined into a larger sentence meaning. In the two-stage account, the meanings of individual words are first used to elaborate sentence meaning, and pragmatic information is then inserted into this meaning. In line with this theoretical distinction between meanings of individual words and pragmatic computations, neuroimaging and neuropsychological studies have demonstrated a specialized neural substrate for the pragmatic comprehension of a speaker's intended meaning, including the medial prefrontal cortex (MPFC), the precuneus, the bilateral posterior superior temporal sulcus (pSTS) and adjacent temporoparietal junctions (for a review, Bara, Enrici, & Adenzato, 2016). Interestingly, while patients with lexico-semantic impairments performed well on pragmatic tasks and were able to communicate their intention, patients with brain damage to their medial prefrontal cortex but with no other linguistic impairments showed deficits in inferring a speaker's intention.

Although speaker identity has been shown to be taken into account in the earliest stages of speech perception (Johnson, Strand, & D'Imperio, 1999; Strand, 1999) and in language comprehension after violations between voices and stereotype-dependent inferences about speakers (Van Berkum et al., 2008), it seems that on-line predictions in language comprehension function in two stages. By highlighting the effects of prediction in two different time windows either integrating prosodic emphasis or not, our findings in Experiment 1 support a two-stage processing account of linguistic predictions. Although an effect of prediction was found only in the first window, suggesting immediate predictions of upcoming words based on sentence meaning driven by prior sentential context, an interactive effect between gender congruency and prosodic emphasis was observed later a second time window. These findings are consistent with those of Corps, Brooke, & Pickering (2022) with gender-stereotyped information related to speakers. This recent eye-tracking study consisting of three different experiments on gender-stereotyped information related to speakers showed that listeners initially predicted language based on the meaning of individual words and then took the speakers' perspective into account (Corps, Brooke, & Pickering, 2022). However, the question arises whether the speaker's perspective was taken into account by listeners in our experiment, as it was in Corps, Brooke, & Pickering (2022)'s study. We studied prosodic emphasis manipulation in the sentential context and an alternative interpretation would be that the difference in prosodic cues with fundamental frequency and mean intensity from the sentential context may lead to late differences in attention without the listener being necessarily aware of the speaker's intention. We did not have an explicit measure to assess whether the listeners inferred the communicative intent of the speaker from the prosodic cues. To provide clearer evidence that the listeners inferred the communicative intent of the speaker, we used a pragmatic judgement task in Experiment 2.

One interesting question is to determine whether pragmatic computations taken into account in spoken-language comprehension are dependent on the endogenous attention to these computations. In Experiment 2, we explored the impact of endogenous attention to the intentional aspect of communication on the neural signatures of predictions in spoken language comprehension.

5. Methods: Experiment 2

Thirty-two native French speakers, with the same characteristics as in Experiment 1, participated in Experiment 2 (mean age = 21.9 years old, range 19–28 years old, standard deviation: 1.7 years old; 5 males). Five additional participants were excluded owing to excessive blinking and movement artifacts. During the sentence listening task, participants were asked to listen to sentences for comprehension and to perform a pragmatic judgement task about the speaker's intention after the listening phase. They had to indicate as quickly and accurately as possible whether they judged the sentence reproachful or not. Ten native speakers (not included in this experiment) had previously performed this task; their judgments were similar and 25% of the sentences were judged reproachful. This ratio was the same in each experimental block and this corresponded to sentences that the speaker had explicitly described as expressing a reproach after producing them. Thirty % of critical sentences conveyed the intentional communication of reproach and were equally balanced by experimental condition. For instance, the following sentential context "*En présentant l'émission, tu t'es trompé dans le texte qui défilait : tu ne l'as pas bien lu(e),*" (in English: "While hosting the program, you made a mistake in the text that was scrolling as you didn't read it properly",) was produced by the speaker as expressing reproach towards the interlocutor. Fifteen % of all filler sentences conveyed the intentional communication of reproach. A question about the speaker's intention was displayed after a black screen and participants had to give their responses (Figure 1A). A black screen with a gray cross then appeared for 2 s. As in Experiment 1, participants performed a lexical recognition task after listening to all the sentences. Stimuli, data acquisition, pre-processing and analyses were the same as in Experiment 1.

6. Results: Experiment 2

6.1 Behavioral results

In the pragmatic judgement task, individual values of *d*-prime differed significantly from the null hypothesis of chance performance (mean value of *d*-prime: 2.03, $t(31)=30$, $p=1.81\times10^{-24}$), suggesting that participants had no difficulty in judging the intentional communication of reproach. An ANOVA analysis on false alarms with Congruency of gender and Prosodic emphasis as factors revealed that participants produced fewer false alarms to judge the speaker's intention when the sentential context was emphasized (EGE+: 0.11, UGE+: 0.12; EGE-: 0.22, UGE-: 0.19, main effect of Prosodic emphasis, $F(1,31)=41.55$, $p=3.47\times10^{-7}$). There were neither a main effect of Congruency of gender nor interaction between Congruency of gender and Prosodic emphasis on false alarms. As in Experiment 1, individual values of *d*-prime of lexical recognition task also differed significantly from the null hypothesis of chance performance in Experiment 2 ($t(31)=14.27$, $p=3.57\times10^{-17}$), showing that participants paid attention to the meaning of sentences (mean value of *d*-prime: 0.36). An ANOVA analysis on false alarms including one factor with five modalities (EGE-, UGE-, EGE+, UGE+, new words) revealed the same main effect between the false alarms of four experimental conditions (EGE-, UGE-, EGE+, UGE+) and those of new words, as in Experiment 1 ($F(4,124)=19.32$,

$p=2.25\times10^{-12}$). By comparing the false alarms of one experimental condition to those of new words, post-hoc Tukey tests showed that participants produced more false alarms for expected yet unheard words than for new words (i.e. unexpected and unheard words, EGE-: 0.44, UGE-: 0.39, EGE+: 0.42, UGE+: 0.40, new words: 0.24, $ps=1.7\times10^{-6}$), suggesting that they paid attention to the meaning of sentential contexts. There were no significant effects between experimental conditions on false alarms.

6.2 EEG results

As in Experiment 1, we compared the TFRs of alpha (8 – 12 Hz), beta (16 – 20 Hz), low gamma (25 – 40 Hz) and high gamma (60 – 100 Hz) band activities for the UG and EG conditions in the interval between 0 and 1s after adjective onset. The cluster-based permutation tests revealed no significant differences in TFRs between the UG and EG conditions at any frequency bands. However, the statistical analysis showed a significant cluster over high gamma (60 – 100 Hz) when comparing the differences in TFRs for UG minus EG between the E- and E+ conditions. As seen in Figure 5, this corresponded to the observed interaction between gender congruency and prosodic emphasis over the frontocentral and centroparietal electrodes in a time window between 0.75 and 0.95 s after adjective onset ($p=4.3\times10^{-2}$). High gamma power was suppressed with respect to the baseline. When the context was produced without emphasis, the unexpected gender elicited a greater decrease in high gamma power than the expected gender ($p=4\times10^{-2}$). This prediction effect was not found when the context was produced with emphasis.

< Insert Figure 5 here >

7. Discussion: Experiment 2

The beneficial effect of emphasized sentential context on the performances in the pragmatic judgement task is in line with the idea that accented speech causes beneficial effects on language processing by capturing more attention (Cutler, 1976; Sanford, Sanford, Molle, & Emmott, 2006). Although the behavioral results of the lexical recognition task were similar to those of Experiment 1 and showed that participants paid attention to the meaning of sentential contexts, the neural signature of predictions in spoken language comprehension was quite different from that of Experiment 1. The endogenous attention to the speaker's intention seemed to reduce the predictive processing in spoken language comprehension. Indeed, the greater decrease in alpha power elicited by the expected gender with respect to the unexpected gender in a time window between 0 and 0.23 s after adjective onset was not observed in Experiment 2. These results thus suggest that the semantic content of sentential context is mainly used to predict upcoming words, and that immediate prediction of upcoming words occurs when listeners focus on sentence meaning. We found another late interactive effect between gender congruency and prosodic emphasis, although it occurred later than in Experiment 1. Unlike in the latter, this late interaction between gender congruency and prosodic emphasis manifested in rapid oscillations at higher frequencies such as the high-gamma band. Interestingly, the endogenous attention to the speaker's intention was so disruptive that prediction appeared only when the context was produced without emphasis. In summary, we found that endogenous attention to pragmatic computations disrupts word prediction, whereas endogenous attention to sentence meaning helps to elicit immediate word prediction. In a later stage, an opposite pattern was found according to the attention focused on pragmatic level or on sentence meaning, with evidence of word prediction when the context was produced with emphasis or without it. These findings therefore support the idea that the meanings of individual words and pragmatic computations are functionally distinct in the processes of language comprehension (Bara, Enrici, & Adenzato, 2016; Clark, 1996; Van Berkum et al., 2008).

Regarding the oscillatory correlates of linguistic predictions, the endogenous attention to the speaker’s intention produced changes in high-gamma activity over a late period. In line with the functional role of the gamma band described as reacting to the matching between top-down predictions and incoming input (see Lewis, & Bastiaansen, 2015; Meyer, 2018, for reviews), high-gamma activity was affected by the congruency of gender between the critical adjectives and the predicted words. High-gamma oscillations have been suggested to reflect the propagation of prediction errors in language comprehension and in speech perception (Arnal, & Giraud, 2012; Lewis, & Bastiaansen, 2015). A reduced gamma activity after the incoming adjective when its gender did not match that of the word predicted from sentential context may however be explained by the attentional manipulation in this experiment. Attention has been suggested to influence gamma activity, indicating that gamma fluctuations are therefore dependent on experimental designs and task strategies (Gruber et al., 1999). Some electrophysiological studies of sentence processing have even found reduced gamma activity with semantically anomalous/less predictable target words (Rommers et al., 2013; Wang, Zhu & Bastiaansen, 2012; Willems et al., 2008, for review see Prystauka, & Lewis, 2019, p. 15) relative to predictable/congruent targets. As what drives the differences in high-gamma activity remains elusive, future studies should address how this activity is affected by endogenous attention during predictions in sentence processing.

8. General Discussion

In the present study, we investigated neural signatures of linguistic predictions and how they are modulated by listener’s attention. We distinguished exogenous from endogenous attention and manipulated the prosodic emphasis produced by the speaker. Clear evidence of immediate word prediction was found when listeners had to focus on the meaning of sentences.

Alpha activity was more strongly reduced during the early processing of an incoming adjective when its gender matched that of the word predicted from sentential context. In later stages, the prosodic emphasis was integrated into word prediction. A greater decrease in alpha activity for matching adjectives was observed only when the context was produced with emphasis. When listeners focused on the pragmatic information, immediate word prediction disappeared and late integration of the prosodic emphasis into word prediction appeared as high-gamma oscillations. In light of the previous literature, the implications of these findings are discussed below.

Alpha oscillations carry predictions in spoken-language comprehension and interplay with listener's attention to prosodic emphasis

Alpha-band oscillations are known to be associated with attentional engagement and successful listening comprehension (Boudewyn, & Carter, 2018). EEG measurements revealed that a greater percentage of alpha oscillations after episodes of mind-wandering during listening comprehension was linked to decreased comprehension accuracy (Boudewyn, & Carter, 2018). In nonlinguistic domains, predicting upcoming information causes attention to focus on the predicted information with a greater decrease in alpha activity after the predicted information (Foxe et al., 1998; Mayer et al., 2016). Some authors have thus suggested that alpha activity reflects the direction of attention from external stimuli for successful performance (Jensen et al., 2002; Jensen, & Mazaheri, 2010; Mazaheri, & Jensen, 2010; Roux, & Uhlhaas, 2014; Strauß et al., 2014; Weisz et al., 2011; Wilsch, & Obleser, 2016). Interestingly, alpha suppression was not higher after a 100% predictable target event (Bidet-Caulet et al., 2012), which shows that reduced alpha oscillations are manifested by attentional preparation of the upcoming information.

Our study provides the first evidence that alpha oscillations carry predictions in spoken-language comprehension with an experimental design disentangling predictive from integrative processes. Electrophysiological studies in language comprehension have already shown that strongly constraining sentences exhibit a decrease in alpha activity related to weakly constraining sentences (Hustá, Zheng, Papoutsi, & Piai, 2021; Lé́on-Cabrera, Piai, Morís, & Rodríguez-Fornells; Piai, Roelofs, & Maris, 2014; Rommers et al., 2017; Roos, & Piai, 2020; Wang et al., 2018), which supports the major role of alpha oscillations in the attentional preparation of upcoming information. Regarding the prediction effect, Molinaro et al. (2017) found lower beta power for gender-mismatching than for gender-matching adjectives in a word-by-word presentation. However, Brunellière et al. (2022) reported a prediction effect over the low-gamma band with auditorily presented adjectives after the audiovisual presentation of sentential context. Although it was not surprising to find alpha activity in the experiment in which we investigated allocating attention triggered by prosodic emphasis, it appeared that the oscillatory correlates of linguistic prediction in language comprehension vary as a function of modality presentation (written vs. spoken language) or endogenous attention to the linguistic information (sentence meaning vs. pragmatic information).

Moreover, alpha oscillations seem to be sensitive to various types of predictions. When listeners focused on the understanding of sentence meaning only, alpha activity was first associated with predictions from the semantic content of sentential context independently of the prosodic emphasis, and then with predictions taking into account the prosodic emphasis from which pragmatic cues may be extracted. These findings are in accordance with a two-stage account of language comprehension (Clark, 1996; Van Berkum et al., 2008), which assumes that sentence meaning is first understood by using the meaning of individual words and then by adding pragmatic information. Another recent eye-tracking study consisting of three different experiments on gender-stereotyped information related to speakers contrasted the time-course of predictions based on the semantic content of sentential context with that of predictions based on the speaker's perspective (Corps, Brooke, & Pickering, 2022). In this study, listeners first predicted language based on the meaning of individual words and then took the speaker's perspective into account. Our findings on alpha oscillations and those by Corps, Brooke, & Pickering (2022) thus support the theory of a two-stage account for linguistic predictions in language comprehension. It can be accounted for by the various levels of linguistic information, including the lexico-semantic and pragmatic information used to produce sentence meaning from the input. This is in line with neuroimaging and neuropsychological studies which have demonstrated a specialized neural substrate for the pragmatic comprehension of a speaker's communication intention, including the medial prefrontal cortex (MPFC), the precuneus, the bilateral posterior superior temporal sulcus (pSTS) and the adjacent temporoparietal junctions (for a review, Bara, Enrici, & Adenzato, 2016). Some authors have also suggested that interpreting speakers' perspectives is an effortful process (e.g. Lin, Keysar, & Epley, 2010). Studies in psychology of dialogue have shown that listeners tend to ignore the perspectives of their interlocutors and adopt an egocentric position (e.g. Keysar, Barr, Balin, & Brauner, 2000). Therefore, pragmatic information takes more time than semantic content to be integrated into linguistic predictions.

In line with the theoretical view assuming that word prediction is optional (Pickering, & Gambi, 2018), our study shows that changes in alpha oscillations related to linguistic predictions seem to be optional in the second step during which the prosodic emphasis is integrated into linguistic predictions. While these changes were observed when the context was produced with emphasis, no linguistic prediction occurred when the context was produced without it. More than showing the allocation of attention to accented words elicited more detailed semantic processing (Birch, & Garnsey, 1995; Blutner, & Sommer, 1988; Fraundorf, Watson, & Benjamin, 2010; Li, & Lu, 2011), we demonstrate that the listener's attention to prosodic emphasis produces changes in alpha oscillations related to linguistic predictions during language comprehension.

High-gamma oscillations support the interplay between predictions in spoken-language comprehension and prosodic emphasis when attention is focused on pragmatic information

In the literature on the predictive coding theory (Bar, 2007; Friston, 2005), high-gamma oscillations are assumed to reflect prediction error signals when the predicted information does not match the incoming information. Prediction error signals are used to update representations and are helpful for learning new information. In language comprehension, high-gamma band power has been found to be higher with semantically incongruent words than with semantically congruent ones (Wang et al., 2018). The propagation of prediction errors was evidenced by high-gamma band power in speech perception (Arnal, & Giraud, 2012; Lewis, & Bastiaansen, 2015). While our study replicated the high-gamma band power reacting to the matching between top-down predictions and incoming input, the pattern of reduced gamma activity after the incoming adjective when its gender mismatched that of the word predicted from sentential context does not support the hypothesis that prediction errors trigger changes in high-gamma oscillations. However, some authors also argued that high-gamma band power reflects a temporal binding phenomenon concerning the coordination of neural assemblies involved in accessing meaning from sensory input (Peña, & Melloni, 2012). For instance, when listeners were exposed to their native language, high-gamma band power decreased after each word until the end of the sentences, owing to a stronger binding effect related to the building of sentence meaning. Previous studies also found greater high-gamma activity in response to highly expected words than to unexpected ones (e.g. Monsalve, Pérez, & Molinaro, 2014; Penolazzi, Angrilli, & Job, 2009). Some electrophysiological studies of sentence processing have found reduced gamma activity with semantically anomalous/less predictable target words (Rommers et al., 2013; Wang, Zhu & Bastiaansen, 2012; Willems et al., 2008, for review see Prystauka, & Lewis, 2019, p. 15) relative to predictable/congruent targets.

In our study, high-gamma oscillations were found to be associated with linguistic predictions when listeners focused their attention on the pragmatic information related to the speaker's intention. However, when listeners focused their attention on sentence meaning, alpha oscillations were the neural signature of linguistic predictions in spoken-language comprehension. These findings are consistent with the idea that high-gamma oscillations were associated with attentional language control depending on pragmatic and social aspects (Tomić, & Kaan, 2022). Moreover, in line with the functional role of high-gamma fluctuations reflecting a binding phenomenon depending on tasks (Fries, 2005; Singer, 1999; Varela, et al., 2001), high-gamma oscillations related to linguistic predictions might emerge especially when the task requires attention on the manipulated information. Attention has already been suggested to influence gamma activity, indicating that gamma fluctuations are dependent on experimental designs and task strategies (Gruber et al., 1999). On the contrary, alpha activity has been associated with the attentional preparation of upcoming information from external cues indirectly related to the task. Beyond the neural signature of linguistic prediction in spoken-language comprehension, the attention focused on pragmatic information reduced the involvement of linguistic predictions and led listeners to produce such predictions integrating prosodic emphasis. Future studies should replicate these findings by manipulating other speech act categories (criticism, doubt, etc.).

Conclusion

We found that the oscillatory correlates of linguistic predictions in language comprehension are highly dependent on the listener's endogenous attention to the level of linguistic information.

While evidence of word prediction has been related to alpha fluctuations when listeners focus their attention on sentence meaning, changes in high-gamma oscillations are triggered by word prediction when they focus their attention on speakers' communication intention. Independently of the endogenous attention to a level of linguistic information, oscillatory correlates of linguistic predictions in language comprehension are sensitive to prosodic emphasis at a late stage. This study demonstrates the high impact of listeners' attention to speakers' communication intention on linguistic predictions in spoken-language comprehension. Researchers can find the raw EEG data online at the following url https://doi.org/10.5281/zenodo.7650679, with scripts of EEG data analyses, false alarms data, and sentence materials.

Figure Captions

Figure 1. Overview of experimental procedure and stimuli. (A) Experimental procedure of sentence listening task; (B) Example of experimental stimuli with gender manipulation. The following sentential context: "*Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi…*" predicted the masculine word "*conseil*". The possessive adjective referring to the second person, "*ton*", is the masculine form and therefore the expected gender in this example. The possessive adjective referring to the second person, "*ta*", is the feminine form and therefore the unexpected gender in this example; (C) Experimental procedure of lexical recognition task. EGE-: Expected gender and without emphasis; UGE-: Unexpected gender and without emphasis; EGE+: Expected gender and with emphasis; UGE+: Unexpected gender and with emphasis.

Figure 2. Spectrograms of examples of sentence frames in each experimental condition. Intensity shown in yellow with scale from 50 to 100 dB. Fundamental frequency (F0) shown in blue with scale from 75 to 630 Hz. Dotted lines correspond to main boundaries in the sentence frames. Expected adjective in green and unexpected adjective in red.

Figure 3. Results based on gender congruency effect on 8 – 12 Hz alpha activity after target adjective in Experiments 1 and 2; (A) Time-frequency representations and power spectral density over the set of significant electrodes for the gender congruency effect from 0 s to 1 s after onset of target adjective in Experiments 1 and 2. Target adjective started at 0 s. Grey bars denote period of significant clusters; (B) Spatial distribution of power elicited by UG and EG conditions over significant time window found in Experiment 1. Asterisks denote significant clusters of electrodes and circles indicate no significant effect of gender congruency at one electrode.

Figure 4. Results based on interactive effect between gender congruency and prosodic emphasis on 8 – 12 Hz alpha activity after target adjective in Experiments 1 and 2; (A) Time-frequency representations and power spectral density over the set of significant electrodes for gender congruency effect at each level of prosodic emphasis from 0 s to 1 s after onset of target adjective in Experiments 1 and 2. Target adjective started at 0 s. Grey bars denote period of significant clusters; (B) Spatial distribution of UG and EG power difference over significant time window found in Experiment 1. Asterisks denote significant clusters of electrodes and circles indicate no significant effect of gender congruency at one electrode.

Figure 5. Results based on interactive effect between gender congruency and prosodic emphasis on 60 – 100 Hz high-gamma activity after target adjective in Experiments 1 and 2; (A) Time-frequency representations and power spectral density over the set of significant electrodes for gender congruency effect at each level of prosodic emphasis from 0 s to 1 s after onset of target adjective in Experiments 1 and 2. Target adjective started at 0 s. Grey bars denote period of significant clusters; (B) Spatial distribution of UG and EG power difference over significant time window found in Experiment 2. Asterisks denote significant clusters of electrodes and circles indicate no significant effect of gender congruency at one electrode.

Supplementary material

Table 1. Psycholinguistic properties for each sub-group of words including in the data analyses of the lexical recognition task.

	Group1	Group2	Group3	Group4	Group5
Lexical frequency (per million occurrences, film subtitles; books)	mean: 35.72, range: 1-232, SD: 56.2; mean: 32.47, range: 1-191, SD: 45.28	mean: 32.98, range: 2-257, SD: 45.18; mean: 28.67, range: 2.4-154.9, SD: 33.4	mean: 38.8, range: 1-391, SD: 60.18; mean: 20.47, range: 1-142.6, SD: 30.6	mean: 31.75, range: 1-250, SD: 40.15; mean: 26.54, range: 1-128, SD:30.36	mean: 30.25, range: 1-220 SD: 35.18; mean: 22.31, range: 1-133.1, SD: 30.37
Number of letters	mean: 7.7, range: 4-8, SD: 1.03	mean: 7.4, range: 4-9, SD: 1.39	mean: 6.75, range: 5-10, SD: 1.41	mean: 7.2, range: 5-10, SD: 1.28	mean: 7.6, range: 5-10 SD: 1.54
Number of phonemes	mean: 5.2, range: 4-7, SD: 0.83	mean: 5.75, range: 4-7, SD: 0.78	mean: 5.9, range: 4-8, SD: 1.21	mean:5.5, range: 4-8, SD: 1	mean: 5.95, range: 4-8, SD: 1.17
Orthographical neighborhood	mean: 2.55, range: 0-16, SD: 2.66	mean: 2, range: 0-10, SD: 1.44	mean: 1.5, range:0-10, SD 1.54:	mean: 1.95, range: 0-12, SD: 2.2:	mean: 1.35, range: 0-7, SD: 1.95
Phonological neighborhood	mean: 6.9, range: 0-29, SD: 8.12	mean: 4.85, range: 0-19, SD: 6.16	mean: 4.1, range:0-15, SD: 5.7	mean: 5, range: 0-17, SD: 6.4	Mean: 5.5, range: 0-20, SD: 4.98

Note. Groups1-4 are groups of words expected from the sentence frames and yet muted, Group5 is a portion of new words; SD: Standard Deviation.

Acknowledgements

This research was supported by a grant for visual studies (SCV2013-2014) from the French National Research Agency (ANR-11-EQPX-0023), European funds through the FEDER SCV-IrDIVE program, and another grant from the French National Research Agency (ANR-19-CE28-0006). It was also funded by the University of Lille (AAPEtablissement2015 & AAPEtablissement2014), the municipal authorities in Lille (AppelLMCU2013), by the *Maison Européenne des Sciences de l'Homme et de la Société* and by the *Conseil Régional des Hauts-de-France* (AAPPartenariats2019). We are very grateful to Chloé Monnier, Emmanuel Farce, Maëlle Legrand, Ombline Nuez, Giulia Di Gregorio and Laurent Ott for their help in selecting the stimuli and running the experiment. We also thank Benjamin Lob for recording the stimuli and three anonymous reviewers for their help during the peer review process. The manuscript was proofread by a native-speaking English copyeditor.

References

Arnal, L.H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*, 390-398. https://doi.org/10.1016/j.tics.2012.05.003

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280-289. https://doi.org/ 10.1016/j.tics.2007.05.005

Bara, B.G., Enrici, I., & Adenzato, M. (2016). The neural substrates of communicative intentions, In *Neurobiology of Language*, Hickok, G., Small S., editors, Academic Press, pp. 675-685.

Bhandari, P., Demberg, V., & Kray, J. (2022). Predictability effects in degraded speech comprehension are reduced as a function of attention. *Language and Cognition*, 1-18, https://doi.org/10.1017/langcog.2022.16

Bidet-Caulet, A., Barbe, P.G., Roux, S., Viswanath, H., Barthélémy, C., Bruneau, N., Knight, R., & Bonnet-Brilhault, F. (2012). Dynamics of anticipatory mechanisms during predictive context processing. *European Journal of Neuroscience*, *36*, 2996-3004. https://doi.org/10.1111/j.1460-9568.2012.08223.x

Birch, S.L., & Garnsey, S.M. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language*, *34*, 232-267. https://doi.org/10.1006/jmla.1995.1011

Blutner, R., & Sommer, R. (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language*, *27*, 359-367. https://doi.org/10.1037/e636952013-638

Boersma, P. & Weenink, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 3.4, retrieved 2 Jan 2011 from http://www.praat.org/

Boudewyn, M.A., & Carter, C.S. (2018). I must have missed that: Alpha-band oscillations track attention to spoken language. *Neuropsychologia*, *117*, 148-155. https://doi.org/10.1016/j.neuropsychologia.2018.05.024

Brunellière, A., Auran, C., & Delrue, L. (2019). Does the prosodic emphasis of sentential context cause deeper lexical-semantic processing? *Language, Cognition & Neuroscience*, *34*, 29-42. https://doi.org/10.1080/23273798.2018.1499945

Brunellière, A., Delrue, L., & Auran, C. (2020). The contribution of audiovisual speech to lexical-semantic processing in natural spoken sentences. *Language, Cognition & Neuroscience*, *35*, 694-711. https://doi.org/10.1080/23273798.2019.1641612

Brunellière, A., & Soto-Faraco, S. (2015). The interplay between semantic and phonological constraints during spoken-word comprehension. *Psychophysiology*, *52*, 46-58. https://doi.org/10.1111/psyp.12285

Brunellière, A., & Soto-Faraco, S. (2013). The speakers' accent shapes the listeners' phonological predictions during speech perception. *Brain and Language*, *125*, 82-93. https://doi.org/10.1016/j.bandl.2013.01.007

Brunellière, A., Vincent, M., & Delrue, L. (2022). Oscillatory correlates of linguistic prediction and modality effects during listening to auditory-only and audiovisual sentences. *International Journal of Psychophysiology*, *178*, 9-21. https://doi.org/10.1016/j.ijpsycho.2022.06.003

Clark, H.H. (1996). *Using language*. Cambridge, MA: Cambridge University Press.

Cohen, L., Salondy, P., Pallier, C., & Dehaene, S. (2021). How does inattention affect written and spoken language processing? *Cortex*, *138*, 212-227. https://doi.org/10.1016/j.cortex.2021.02.007

Corps, R.E, Brooke, C. &, Pickering, M.J. (2022). Prediction involves two stages: Evidence from visual-world eye-tracking. *Journal of Memory and Language*, *122*, 104298, https://doi.org/10.1016/j.jml.2021.104298

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics*, *20*, 55-60. https://doi.org/10.3758/bf03198706

DeCarlo, L.T. (1998). Signal detection theory and generalized linear models. *Psychological Methods*, *3*(2), 186-205. https://doi.org/10.1037/1082-989X.3.2.186

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117-1121. https://doi.org/10.1038/nn1504

Fleur, D.S., Flecken, M., Rommers, J., & Nieuwland, M.S. (2020). Definitely saw it coming? The dual nature of the pre-nominal prediction effect. *Cognition*, *204*, e104335. https://doi.org/10.1016/j.cognition.2020.104335

Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75-84. doi:10.1016/j.brainres.2006.06.101

Foucart, A., Martin, C.D., Moreno, E.M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461-1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). How do you know I was about to say "book"? Anticipation processes affect speech processing and lexical recognition. *Language, Cognition and Neuroscience*, *30*, 768-780.

https://doi.org/10.1080/23273798.2015.1016047

Foxe, J.J., Simpson, G.V., & Ahlfors, S.P. (1998). Parieto-occipital–10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport, 9*(17), 3929-3933. https://doi.org/10.1097/00001756-199812010-00030

Fraundorf, S.H., Watson, D.G., & Benjamin, A.S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, *63*, 367-386. https://doi.org/10.1016/j.jml.2010.06.004

Fridriksson, J., Moss, J., Davis, B., Baylis, G.C., Bonilha, L., & Rorden, C. (2008). Motor speech perception modulates the cortical language areas. *NeuroImage*, *41*, 605-613. https://doi.org/10.1016/j.neuroimage.2008.02.046

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*, 474-480. 10.1016/j.tics.2005.08.011

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the royal society B*, *360*, 815-836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, *22*, 1093-1104. 10.1016/j.neunet.2009.07.023

Gruber, T., Müller, M.M., Keil, A., & Elbert, T. (1999). Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clinical Neurophysiology*, *110*, 2074-2085. https://doi.org/10.1016/s1388-2457(99)00176-5

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, *1*, S18-29. https://doi.org/10.1016/j.neuroimage.2003.09.013

Hald, L. A., Bastiaansen, M.C.M., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language*, *96*(1), 90-105. https://doi.org/10.1016/j.bandl.2005.06.007

Hustá, C., Zheng, X., Papoutsi, C., & Piai, V. (2021). Electrophysiological signatures of conceptual and lexical retrieval from semantic memory. *Neuropsychologia*, *161*, 107988. https://doi.org/10.1016/j.neuropsychologia.2021.107988.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J.E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*, 877-882. http://dx.doi.org/10.1093/cercor/12.8.877.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*. http://dx.doi.org/10.3389/fnhum.2010.00186.

Johnson, K.A., Strand, E.A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *24*, 359-384. https://doi.org/10.1006/jpho.1999.0100

Jun, S.A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, *14*, 147-172. https://doi.org/10.1515/prbs.2002.002

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*, 32-38. https://doi.org/10.1111/1467-9280.00211

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

Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*, 203-204. https://doi.org/10.1126/science.7350657

Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. Paris: CNRS Editions.

León-Cabrera, P., Piai, V., Morís, J., & Rodríguez-Fornells, A. (2022). Alpha power decreases associated with prediction in written and spoken sentence comprehension. *Neuropsychologia*, *173*, 108286. https://doi.org/10.1016/j.neuropsychologia.2022.108286

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126-1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lewis, A.G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex, 68*, 155-168. https://doi.org/10.1016/j.cortex.2015.02.014

Li, X.-Q., & Lu, Y. (2011). How accentuation influences semantic short-term memory representations during on-line speech processing: An event-related potential study. *Neuroscience*, *193*, 217-228. https://doi.org/10.1016/j.neuroscience.2011.07.010

Li, X., Zhang, Y., Xia, J., & Swaab, T.Y. (2017). Internal mechanisms underlying anticipatory language processing: evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, *102*, 70-81. http://dx.doi.org/10.1016/j.neuropsychologia.2017.05.017

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, *46*, 551-556. https://doi.org/10.1016/j.jesp.2009.12.019

Maris, E. & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177-190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mayer, A., Schwiedrzik, C.M., Wibral, M., Singer, W., & Melloni, L. (2016). Expecting to see a letter: Alpha oscillations as carriers of top–down sensory predictions. *Cerebral Cortex, 26*, 3146-3160. https://doi.org/10.1093/cercor/bhv146

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*, 2609-2621. https://doi.org/10.1111/ejn.13748

Mazaheri, A., & Jensen, O. (2010). Rhythmic pulsing: linking ongoing brain activity with evoked responses. *Frontiers in Human Neuroscience*, *4*. http://dx.doi.org/10.3389/fnhum.2010.00177

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition*, *164*, 61-73. https://doi.org/10.1016/j.cognition.2017.03.012

Monsalve, I. F., Pérez, A., & Molinaro, N. (2014). Item parameters dissociate between expectation formats: A regression analysis of time-frequency decomposed EEG data. *Frontiers in Psychology*, *5*, 847. https://doi.org/10.3389/fpsyg.2014.00847

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*, 516-524. https://doi.org/10.3758/BF03195598

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex*, *20*(3), 633-640. https://doi.org/10.1093/cercor/bhp128

Obleser, J., Wise, R. J. S., Dresner, A.M., & Scott, S.K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, *27*(9), 2283-2289. https://doi.org/10.1523/JNEUROSCI.4663-06.2007

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97-113. https://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Otten, M., Nieuwland, M.S., & Van Berkum, J.A. (2007). Great expectations: specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, *8*, 89. https://doi.org/10.1186/1471-2202-8-89

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole* (Doctoral dissertation). Retrieved from A.N.R.T

Peña, M., & Melloni, L. (2012). Brain Oscillations during Spoken Sentence Processing. *Journal of Cognitive Neuroscience*. *24*, 1149-1164. 10.1162/jocn_a_00144

Penolazzi, B., Angrilli, A., & Job, R. (2009). Gamma EEG activity induced by semantic violation during sentence reading. *Neuroscience Letters*, *465*(1), 74-78. https://doi.org/10.1016/j.neulet.2009.08.065

Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, *53*, 146-156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014

Piai, V., Roelofs, A., Rommers, J., & Maris, E. (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping*, *36*(7), 2767-2780. https://doi.org/10.1002/hbm.22806

Pickering, M.J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*, 1002-1044. https://doi.org/10.1037/bul0000158

Posner, M.I., & Cohen, Y. (1984) Components of Visual Orienting. In: Bouma, H. and Bouwhuis, D.G., Eds., *Attention and Performance X: Control of Language Processes*, Erlbaum, Hillsdale, 531-556.

Prystauka, Y., & Lewis, A.G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, *13*, e12347. https://doi.org/10.1111/lnc3.12347

Rommers, J., Dijkstra, T., & Bastiaansen, M.C.M. (2013). Context-dependent semantic processing in the human brain: evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*, 762-776. https://doi.org/10.1162/jocn_a_00337

Rommers, J., Dickson, D.S., Norton, J.J., Wlotko, E.W., & Federmeier, K.D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, *32*(5), 576-589. https://doi.org/10.1080/23273798.2016.1183799

Rommers, J., & Federmeier, K.D. (2018). Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. *Cortex*, *101*, 16-30. https://doi.org/10.1016/j.cortex.2017.12.018

Roos, N.M., & Piai, V. (2020). Across-session consistency of context-driven language processing: A magnetoencephalography study. *European Journal of Neuroscience*, *52*(5), 3457-3469. https://doi.org/10.1111/ejn.14785

Roux, F., & Uhlhaas, P.J. (2014). Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information? *Trends in Cognitive Sciences*, *18*, 16-25. http://dx.doi.org/10.1016/j.tics.2013.10.010

Sanford, A.J.S., Sanford, A.J., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, *42*, 109-130. https://doi.org/10.1207/s15326950dp4202_2

Schneider, T.R., Debener, S., Oostenveld, R., & Engel, A.K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage*, *42*, 1244-1254. https://doi.org/10.1016/j.neuroimage.2008.05.033

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B.A. (2008). Effect of age, presentation method, and learning on identification of noise-vocoded words. *The Journal of the Acoustical Society of America*, *123*(1), 476-488. https://doi.org/10.1121/1.2805676

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, 24, 49-65. https://doi.org/10.1016/s0896-6273(00)80821-1

Strand, E.A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, *18*, 86-99. https://doi.org/10.1177/0261927X99018001006

Strauß, A., Wöstmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, *8*, 350. https://doi.org/10.3389/fnhum.2014.00350

Strube, A., Rose, M., Fazeli, S., & Büchel, C. (2021). Alpha-to-beta- and gamma-band activity reflect predictive coding in affective visual processing. *Scientific reports*, *11*, 23492. https://doi.org/10.1038/s41598-021-02939-z

Tomić, A., & Kaan, E. (2022). Oscillatory brain responses to processing code-switches in the presence of others. *Brain and Language*, *231*, 105-139. https://doi.org/10.1016/j.bandl.2022.105139

Touati, P. (1987). *Structure prosodique du suédois et du français* (Doctoral dissertation). Retrieved from http://www.bibliotheque.bordeaux.fr

Van Berkum, J.J., Brown, C.M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443-467.

Van Berkum, J.J.A., van den Brink, D., Tesink, C.M.J.Y, Kos, M., & Hagoort, P. (2008). The Neural Integration of Speaker and Message. *Journal of Cognitive Neuroscience*, *20*, 580-591. https://doi.org/10.1162/jocn.2008.20054

Van den Brink, D., Brown, C.M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13,* 967-985. https://doi.org/10.1162/089892901753165872

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*, 1068-1084. https://doi.org/10.1162/0898929041502670

Varela, F., Lachaux, J.P., Rodriguez, E., & Martinerie, J. (2001). The brain-web: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2,* 229-239, https://doi.org/10.1038/35067550

Wang, L., Hagoort, P., & Jensen, O. (2018). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Journal of Cognitive Neuroscience*, *30*, 432-447. https://doi.org/10.1162/jocn_a_01190

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language*, *85*(2), 325–343. https://doi.org/10.1016/S0093-934X(03)00067-1

Weisz, N., Hartmann, T., Müller, N., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in Psychology*, *2*, 73. https://doi.org/10.3389/fpsyg.2011.00073

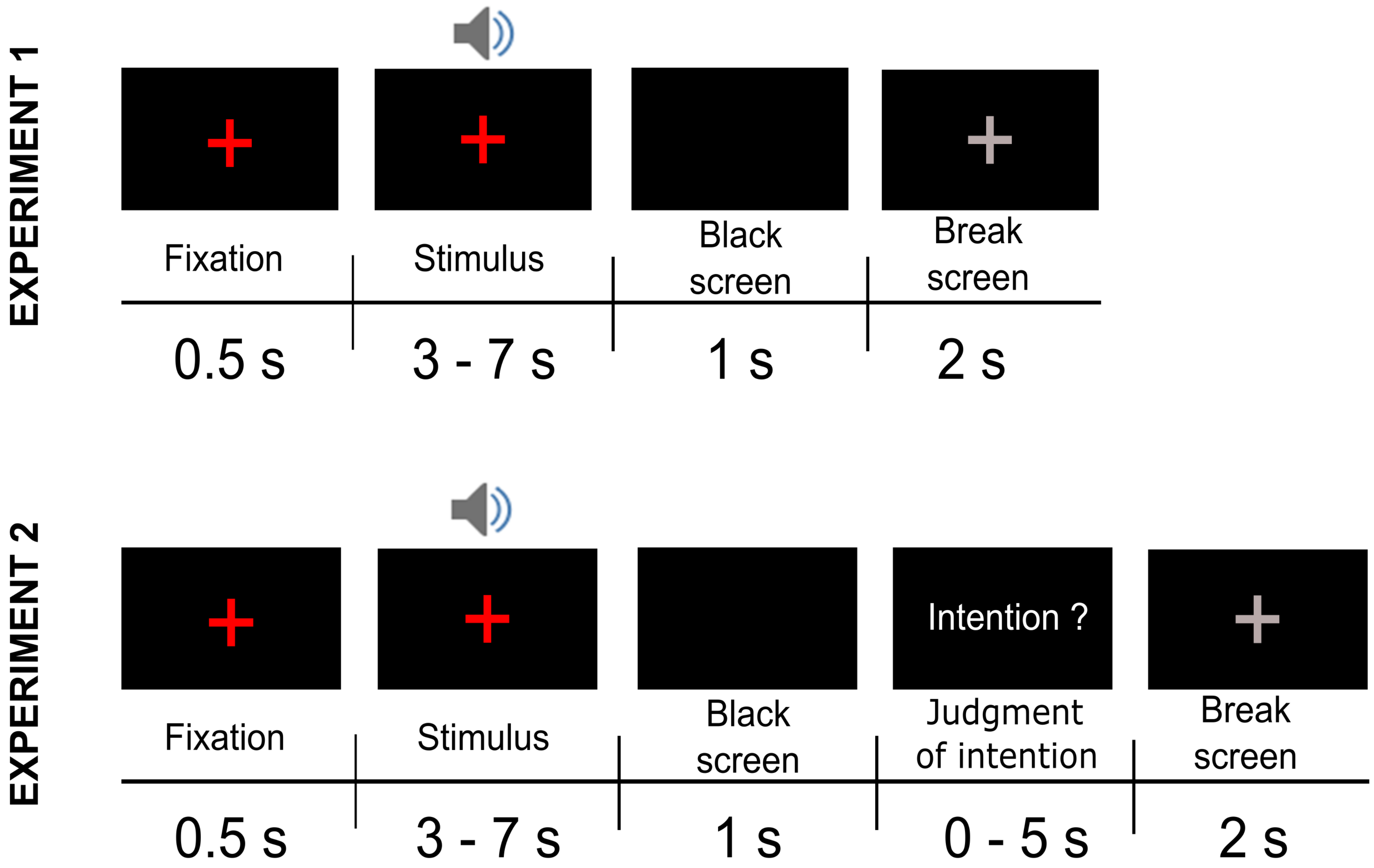
Wicha, N.Y.Y, Bates, E.A, Moreno, E.M., & Kutas, M. (2003). Potato not pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, *346*, 165-168. https://doi.org/10.1016/S0304-3940(03)00599-8

Wicha, N.Y., Moreno, E.M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-

1288. https://doi.org/10.1162/0898929041920487

Wilsch, A., & Obleser, J. (2016). What works in auditory working memory? A neural oscillations perspective. *Brain Research*, *1640*, 193-207. https://doi.org/10.1016/j.brainres.2015.10.054

A. LISTENING TASK



4 CONDITIONS

EGE- : Expected gender and without emphasis

UGE- : Unexpected gender and without emphasis

EGE+ : Expected gender and with emphasis

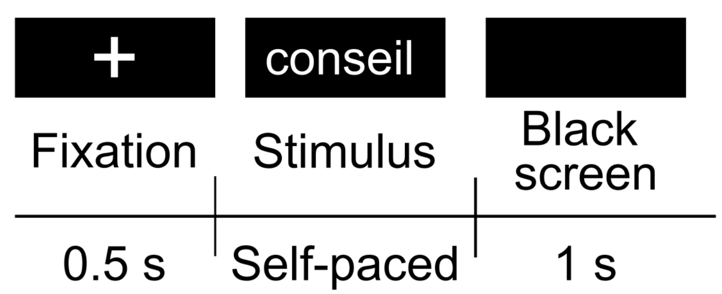
UGE+ : Unexpected gender and with emphasis

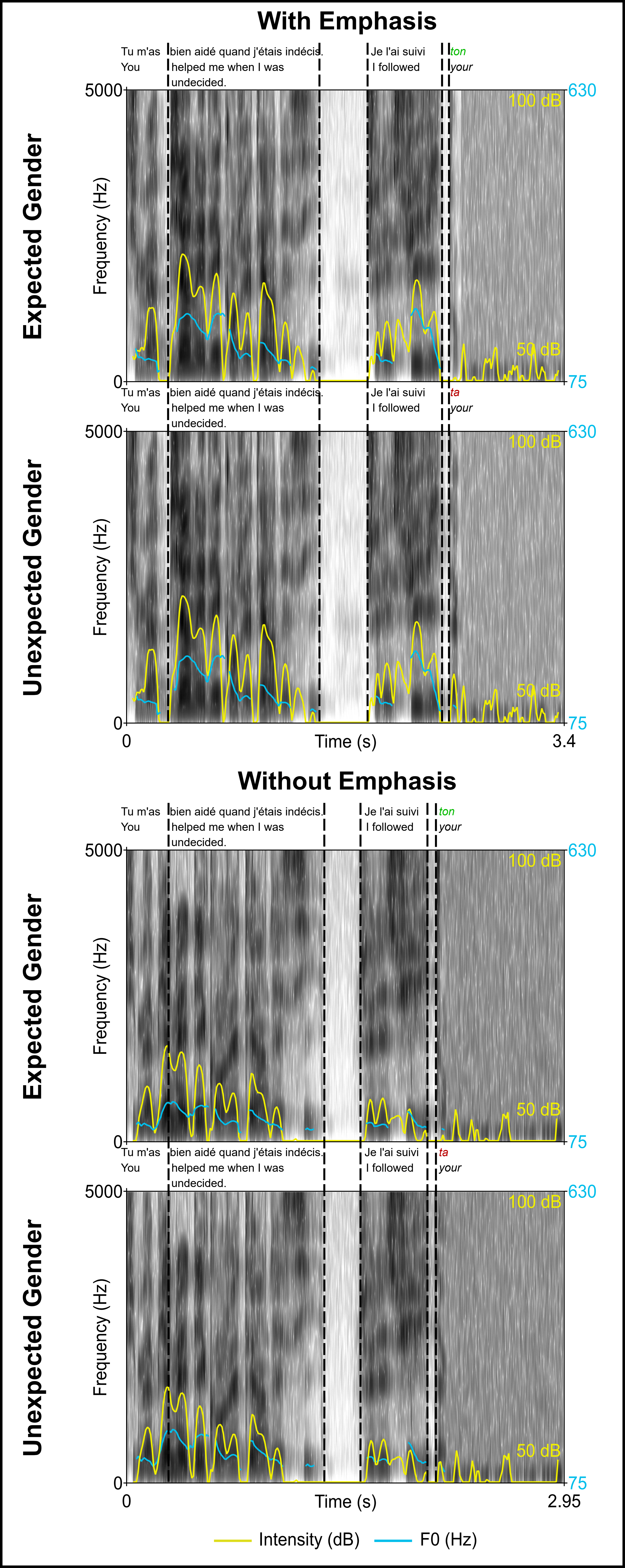
B. EXAMPLE OF SENTENCES WITH GENDER MANIPULATION

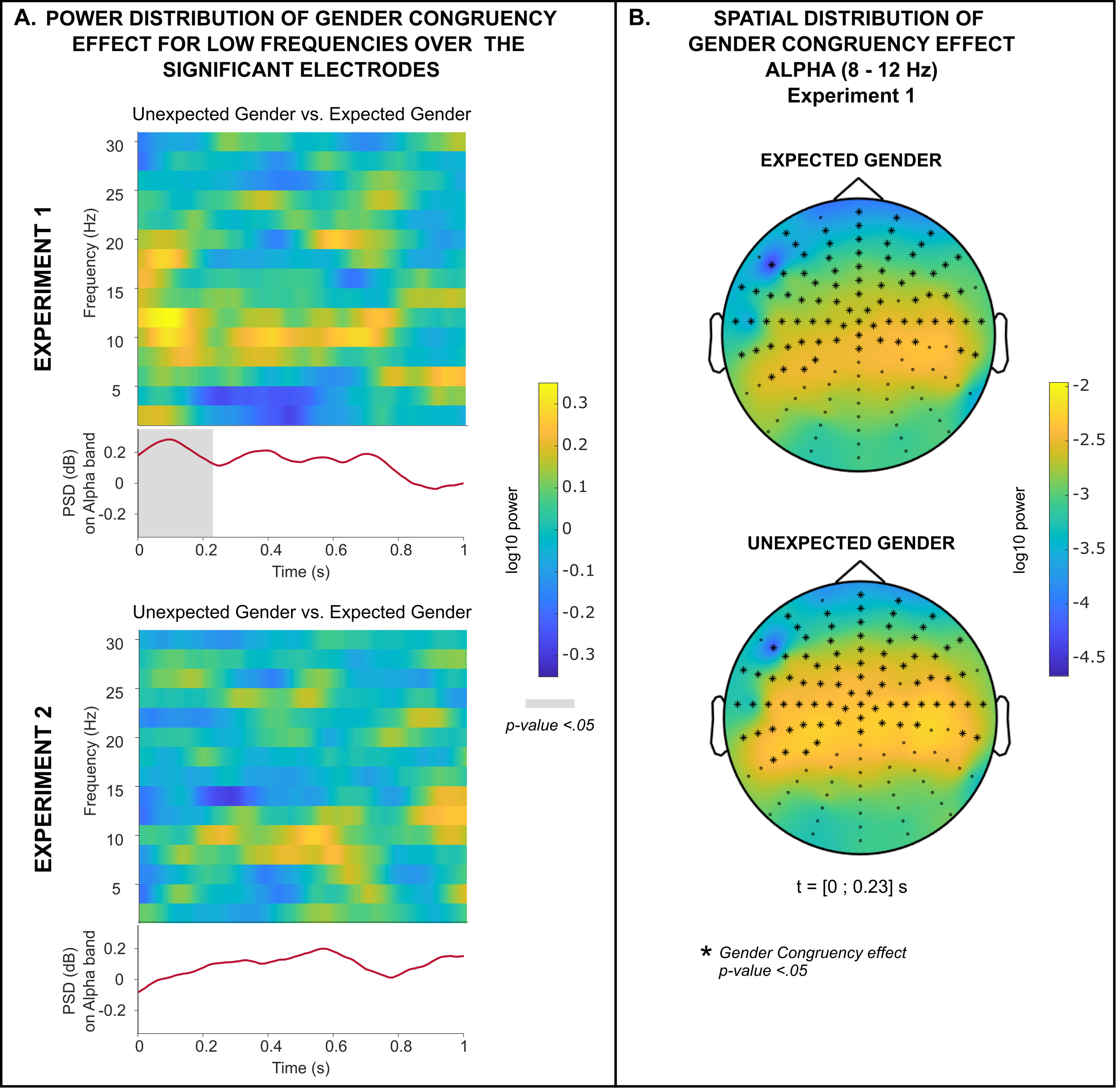
Condition	**Sentence Context**	**Target Adjective**
EG	Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi *You helped me when I was undecided. I followed*	ton *your*
UG	Tu m'as bien aidé quand j'étais indécis. Je l'ai suivi *You helped me when I was undecided. I followed*	ta *your*

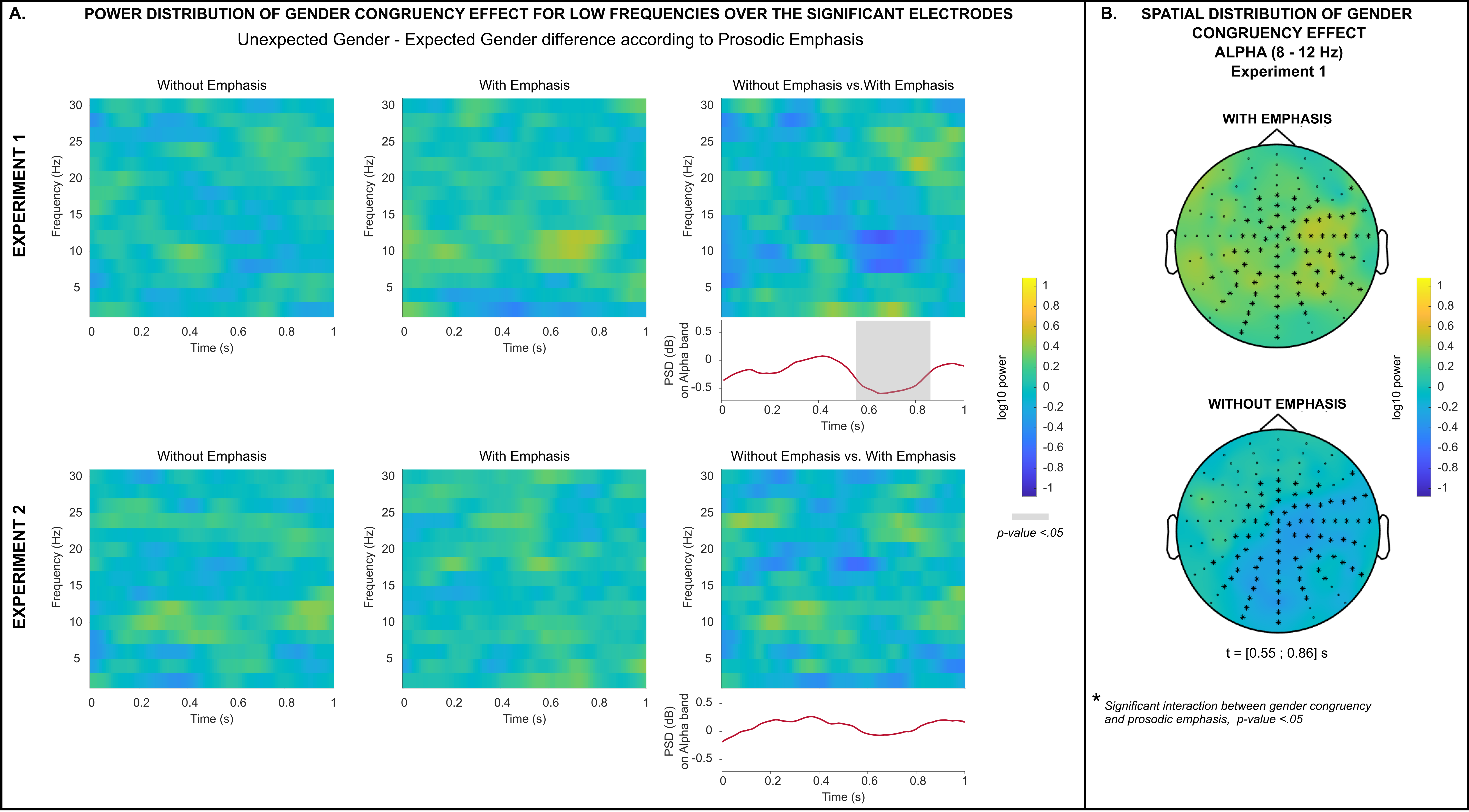
C. LEXICAL RECOGNITION TASK

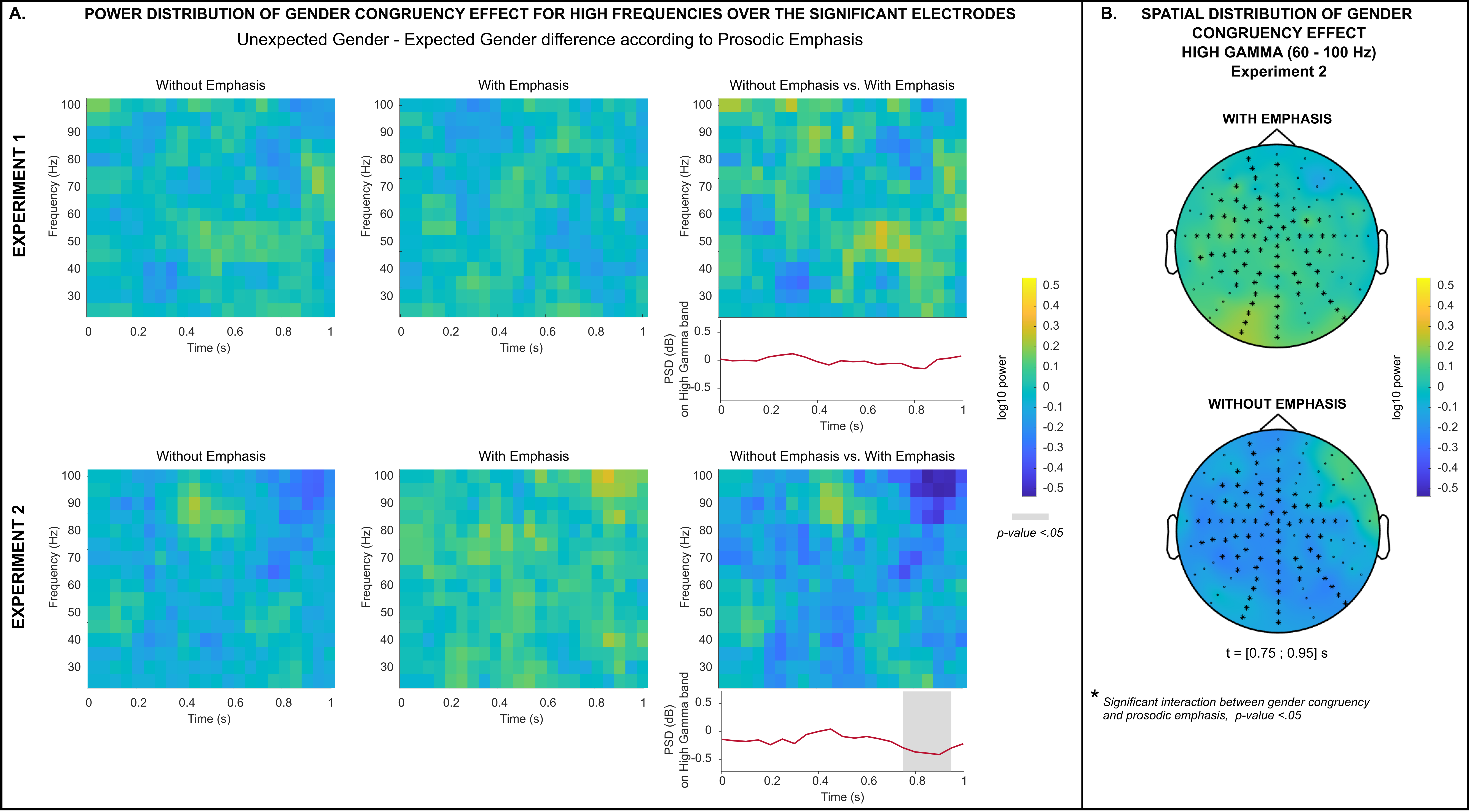
EXPERIMENTS 1 AND 2




With Emphasis
Expected Gender
Tu m'as
You
bien aidé quand j'étais indécis.
helped me when I was
undecided.
Je l'ai suivi
I followed
ton
your
Frequency (Hz)
5000
0
100 dB
50 dB
630
75
Unexpected Gender
Tu m'as
You
bien aidé quand j'étais indécis.
helped me when I was
undecided.
Je l'ai suivi
I followed
ta
your
Frequency (Hz)
5000
0
100 dB
50 dB
630
75
Time (s)
0
3.4
Without Emphasis
Expected Gender
Tu m'as
You
bien aidé quand j'étais indécis.
helped me when I was
undecided.
Je l'ai suivi
I followed
ton
your
Frequency (Hz)
5000
0
100 dB
50 dB
630
75
Unexpected Gender
Tu m'as
You
bien aidé quand j'étais indécis.
helped me when I was
undecided.
Je l'ai suivi
I followed
ta
your
Frequency (Hz)
5000
0
100 dB
50 dB
630
75
Time (s)
0
2.95
Intensity (dB)
F0 (Hz)


A. POWER DISTRIBUTION OF GENDER CONGRUENCY EFFECT FOR LOW FREQUENCIES OVER THE SIGNIFICANT ELECTRODES
EXPERIMENT 1
Unexpected Gender vs. Expected Gender
Frequency (Hz)
PSD (dB) on Alpha band
Time (s)
EXPERIMENT 2
Unexpected Gender vs. Expected Gender
Frequency (Hz)
PSD (dB) on Alpha band
Time (s)
log10 power
p-value <.05
B. SPATIAL DISTRIBUTION OF GENDER CONGRUENCY EFFECT ALPHA (8 - 12 Hz) Experiment 1
EXPECTED GENDER
UNEXPECTED GENDER
log10 power
t = [0 ; 0.23] s
* Gender Congruency effect p-value <.05


A.
POWER DISTRIBUTION OF GENDER CONGRUENCY EFFECT FOR LOW FREQUENCIES OVER THE SIGNIFICANT ELECTRODES
Unexpected Gender - Expected Gender difference according to Prosodic Emphasis
EXPERIMENT 1
Without Emphasis
With Emphasis
Without Emphasis vs.With Emphasis
Frequency (Hz)
Time (s)
PSD (dB) on Alpha band
log10 power
p-value <.05
EXPERIMENT 2
Without Emphasis
With Emphasis
Without Emphasis vs. With Emphasis
B.
SPATIAL DISTRIBUTION OF GENDER CONGRUENCY EFFECT
ALPHA (8 - 12 Hz)
Experiment 1
WITH EMPHASIS
WITHOUT EMPHASIS
log10 power
t = [0.55 ; 0.86] s
* Significant interaction between gender congruency and prosodic emphasis, p-value <.05



A.
POWER DISTRIBUTION OF GENDER CONGRUENCY EFFECT FOR HIGH FREQUENCIES OVER THE SIGNIFICANT ELECTRODES
Unexpected Gender - Expected Gender difference according to Prosodic Emphasis
EXPERIMENT 1
EXPERIMENT 2
Without Emphasis
With Emphasis
Without Emphasis vs. With Emphasis
Frequency (Hz)
Time (s)
PSD (dB) on High Gamma band
log10 power
p-value <.05
B.
SPATIAL DISTRIBUTION OF GENDER CONGRUENCY EFFECT
HIGH GAMMA (60 - 100 Hz)
Experiment 2
WITH EMPHASIS
WITHOUT EMPHASIS
t = [0.75 ; 0.95] s
* Significant interaction between gender congruency and prosodic emphasis, p-value <.05