

# Agreement attraction during comprehension of grammatical sentences: ERP evidence from ellipsis



Andrea E. Martin<sup>a,b,\*</sup>, Mante S. Nieuwland<sup>a,b</sup>, Manuel Carreiras<sup>b,c</sup>

<sup>a</sup> Department of Psychology, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom

<sup>b</sup> Basque Center on Cognition, Brain, and Language, Donostia-San Sebastián, Spain

<sup>c</sup> Ikerbasque Foundation, Basque Foundation for Science, Bilbao, Spain

## ARTICLE INFO

### Article history:

Accepted 1 May 2014

Available online 7 June 2014

### Keywords:

Sentence processing

Memory

Interference

Retrieval cues

ERPs

Anterior negativity

NRef

P600

Ellipsis

Agreement attraction

## ABSTRACT

Successful dependency resolution during language comprehension relies on accessing certain representations in memory, and *not* others. We recently reported event-related potential (ERP) evidence that syntactically unavailable, intervening attractor-nouns interfered during comprehension of Spanish noun-phrase ellipsis (the determiner *otra/otro*): grammatically correct determiners that mismatched the gender of attractor-nouns elicited a sustained negativity as also observed for incorrect determiners (Martin, Nieuwland, & Carreiras, 2012). The current study sought to extend this novel finding in sentences containing object-extracted relative clauses, where the antecedent may be less prominent. Whereas correct determiners that matched the gender of attractor-nouns now elicited an early anterior negativity as also observed for mismatching determiners, the previously reported interaction pattern was replicated in P600 responses to subsequent words. Our results suggest that structural and gender information is simultaneously taken into account, providing further evidence for retrieval interference during comprehension of grammatical sentences.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

Understanding language often requires remembering what was previously said. In the narrowest sense of this ‘remembering’, we access or reactivate linguistic representations that we recently processed in order to form ongoing, coherent structures and integrate incoming information. In other words, we must link linguistic representations through time in a grammatical way, forming linguistic dependencies, and do so often in the face of processing other words or phrases within the dependency:

- (a) Marta se compró la camiseta y Miren cogió otra \_.  
Marta bought the *t-shirt* and Miren took another \_.

To understand this basic Spanish sentence, the noun *t-shirt* (*camiseta*) must be interpreted upon encountering *another* (*otra*) to mean *another t-shirt*. Thus, forming the linguistic dependency between *another* and its antecedent *t-shirt* requires, at minimum, retrieving *t-shirt* from memory. However, at the onset of *another*,

\* Corresponding author at: Department of Psychology, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom.

E-mail address: [andrea.martin@ed.ac.uk](mailto:andrea.martin@ed.ac.uk) (A.E. Martin).

there are presumably other representations in memory that may interfere with the retrieval and interpretation of *t-shirt*, a phenomenon known as *retrieval interference* (Anderson & Neely, 1996; Martin, Nieuwland, & Carreiras, 2012; McElree, 2006; McElree, Foraker, & Dyer, 2003). The often-minimal information that must be used to access the relevant information in memory, or *retrieval cues*, becomes vital to interpretation. What information counts as a retrieval cue in a sentence context is still largely unknown, but an increasing body of evidence implicates syntactic and morphosyntactic, as well as semantic information as giving rise to retrieval interference during dependency formation (Martin et al., 2012; McElree et al., 2003; Van Dyke & McElree, 2011). Because presence of these types of representations in memory appears to create processing difficulty or interference, these findings have been taken as evidence that these types of information may be used to elicit target representations from memory (i.e., function as retrieval cues). Although different levels of linguistic structure have not been extensively tested, it stands to reason that an efficient system would use all relevant types of linguistic representations available at retrieval. Discovering which particular representations control or contribute most to processing of different dependency types may yet inform psycholinguistic and neurobiological theories about the computational architecture of the human language faculty. One example paradigm that has recently informed this

debate is the existence of agreement attraction phenomena during comprehension of subject–verb dependencies such as *the key to the cabinets was* (e.g., Bock & Miller, 1991; Cunnings & Felser, 2013; Dillon, Mishler, Sloggett, & Phillips, 2013; Nicol, Forster, & Veres, 1997; Pearlmutter, Garnsey, & Bock, 1999; Staub, 2009; Tanner, Nicol, Herschensohn, & Osterhout, 2012; Wagers, Lau, & Phillips, 2009). We relate the current findings to existing literature on agreement attraction in comprehension, which generally deals with linguistic phenomena other than ellipsis and (purely) grammatical gender, in our Discussion section.

### 1.1. Language specific cues

In languages like Spanish, words in certain syntactic relationships with one another must agree in grammatical gender. Thus, recovery of the antecedent is likely to be guided by gender cues at retrieval that point to the relevant memory representations (e.g., feminine gender-marking on *otra* that agrees with the gender of the antecedent *camiseta*). However, such a system also entails the possibility that these cues are inadvertently mapped onto other, perhaps more recent or similar, entries in memory:

- (b) Marta se compró la camiseta que iba bien con la falda y Miren cogió otra \_.  
*Marta bought the t-shirt that went well with the skirt and Miren took another \_.*

Here, the same linguistic dependency as in (a) needs to be resolved: *another* must be understood as *another t-shirt*, not as *another skirt*. In English, and in Spanish, this occurs because *skirt* is structurally unavailable<sup>1</sup> as antecedent; structural position alone inherently prevents *skirt* to be the antecedent for *another*. However, in Spanish the structurally unavailable intervening noun (i.e., the local attractor *falda*) also has a grammatical gender that could correspond, as in this example, or not, with *otra/another*. Here the determiner *otra* can only mean *otra camiseta*, and introduces a new discourse referent of the same type as *camiseta*. Importantly, *otra* does not refer to the same instance of *camiseta* (Eguren, 2010), and is therefore not a pronoun or anaphor (which, except for cataphoric pronouns, refer back to a given referent). Eguren (2010) describes *otra* in this instance as a determiner that allows nominal ellipsis (see Martin et al., 2012 for discussion). This determiner must agree with the gender of the elided noun and morphological gender information could thus serve in resolving the dependency (Badecker & Straub, 2002; Sturt, 2003).

### 1.2. Cue-based retrieval interference during processing of ellipsis

The linguistic construction *otro/a* in (a/b) is representative of the type of challenge language users routinely face, establishing basic sentential relations despite a missing argument noun phrase, or interpreting “elided” information. Importantly, ellipsis antecedents bear no special marking that might engage special memory strategies. Behavioral studies of ellipsis processing suggest that interpreting more material between the antecedent and the ellipsis site decreases the likelihood of successful retrieval, and therefore of successful interpretation (Martin & McElree, 2008, 2009, 2011), consistent with an account of retrieval during sentence comprehension where match between cue and target shape processing (Martin et al., 2012; McElree et al., 2003; Nairne, 2002; Van Dyke & McElree, 2011). For the purpose of testing the role of cue-based retrieval interference during language comprehension,

Martin, Nieuwland, & Carreiras (2012; henceforth ‘MNC’) constructed an ellipsis paradigm in Spanish based on (b), but resolved with a determiner that does not agree with the antecedent, ‘otro’ in place of ‘otra’ in (b), creating an ungrammatical condition:

- (c) \*Marta se compró la camiseta que iba bien con el vestido y Miren cogió otro \_.  
 \*Marta bought the t-shirt that went well with the dress and Miren took another \_.

MNC found a sustained negativity to incorrect agreement (c) compared to correct agreement (b), and this negativity was modulated by the gender of a local attractor noun (either mismatching attractor *vestido* or matching attractor *falda*). Grammatically correct determiners following a locally mismatching noun (*camiseta...vestido...otra*) elicited a reliably more negative-going ERP response than following a locally matching attractor noun (*camiseta...falda...otra*). This pattern suggests that an illicit attractor that does not match the retrieval cue impacts or disrupts processing, even in a fully grammatical sentence. The interaction suggests that the attractor gender affects antecedent retrieval, but that similarity or feature overlap alone does not drive the effect. The sustained negativity elicited by incorrect determiners was surprising because previous literature has consistently reported that local agreement violations (e.g., *el camiseta*: Barber & Carreiras, 2005; see also Hanulíková, Van Alphen, Van Goch, & Weber, 2012) elicit P600 effects, a broadly distributed positive shift that is most commonly associated with syntactic computations (e.g., Hagoort, Brown, & Osterhout, 1999). P600 effects are not only reliably elicited by syntactic violations (e.g., Barber & Carreiras, 2005; Osterhout & Mobley, 1995), but also by constructions that are well-formed but whose syntactic properties do not fit the analysis currently being pursued or that was previously expected (e.g., Kaan & Swaab, 2003; Osterhout & Holcomb, 1992; Van Berkum, Brown, & Hagoort, 1999). Notably, pronouns with unexpected or incorrect gender-marking also evoke P600 effects (e.g., Nieuwland & Van Berkum, 2006; Osterhout & Mobley, 1995; Silva-Pereyra, Gutierrez-Sigut, & Carreiras, 2012; Van Berkum, Koornneef, Otten, & Nieuwland, 2007). The lack of a P600 effect reported by MNC, however, can be directly linked to previous ERP findings on the resolution of sentence- and discourse-level referential dependencies. Sustained negativities (the Nref; Van Berkum et al., 1999; see Nieuwland & Van Berkum, 2008a, for review) have been associated with referentially problematic expressions, e.g., referential ambiguity arising from having multiple suitable candidate antecedents in the discourse model. In light of these results, although an elided constituent is not a pro-form or a trace (Eguren, 2010), grammatically incorrect determiners during ellipsis might be problematic for referential reasons (i.e., insufficient to retrieve a unique antecedent). Such is also assumed in memory-based processing models of anaphor resolution (e.g., Gerrig & O’Brien, 2005): antecedents that have sufficient features in common with anaphors are automatically elicited from memory, whereas antecedents that do not require additional retrieval processes. Sustained negativities associated with referentially problematic expressions might be an electrophysiological correlate of these additional retrieval processes (Nieuwland & Van Berkum, 2008a,b; although see Burkhardt, 2006, for different ERP results on introducing referents through anaphoric bridging inferences; see also Kaan, Dallas, & Barkley, 2007), possibly related to sustained frontal shifts that are evoked by linguistically complex structures (King & Kutas, 1995).

Given the novelty of the attractor-modulated sustained negativity found by MNC, the current study attempts to expand this finding to sentences with different syntactic relationships between retrieval cues, antecedents and attractors. We used the same

<sup>1</sup> Because *dress* is in a relative clause, and nouns cannot be bound outside of the relative clause they occur in (Chomsky, 1981).

interference paradigm modulo a change in the syntactic structure of the intervening material – an object-extracted relative clause instead of the subject-extracted relative clause used in MCN (see Example b and Table 1). In MNC, antecedents were objects in the matrix clause and subjects of the relative clause. Thus, in MNC, the antecedent was, at different points in the sentence, assigned both object *and* subject role, whereas the attractor was never a subject, perhaps diminishing its distinctness from the antecedent in memory by decreasing its prominence (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009). In the current study, the antecedent is *only* ever an object, it is never a subject. By changing the relative clause to an object-extracted structure in the current study, we changed the syntactic information on the antecedent at the point of retrieval (see Table 1). Divergent syntactic role assignments might increase prominence or distinctiveness between antecedent and attractor in memory, with syntactic role being another dimension of syntactic information that has been shown to affect encoding and retrieval (although under different linguistic dependencies, see Van Dyke, 2007; Van Dyke & Lewis, 2003). As in the paradigm used by MNC, intervening attractors (the feminine *falda* “skirt” in (1) and the masculine *vestido* “dress” in (2)) have recently been processed when *otra* is encountered.

The gender of these attractors matches the feminine *otra* in gender (*falda* in 1) or not (*vestido* in 2). In either case, *otra* refers to *camiseta*, because the attractor is not structurally licensed as antecedent. The gender of the attractor noun is thus irrelevant if the parsing process only takes into account the correct antecedent. However, the mismatching gender of the attractor noun in (2) might nevertheless incur processing costs in this grammatically correct sentence, in which we would expect a similar pattern of ERP results as observed by MNC. Object-extracted relative clauses in Spanish can feature a resumptive pronoun (in Table 1, the *la* in *con la que*; Suñer, 1998). The pronoun agrees with the antecedent and is another instance of the antecedent’s grammatical gender in the sentence – however, only 12 of our 120 items feature a resumptive pronoun. It is possible that changing the syntactic role of the antecedent changes the electrophysiological response profile of the interference effect, such that a typical P600 modulation is elicited. Nevertheless, our prediction that the attractor noun would create interference, leading to a difference even between the two grammatical conditions, remained the same.

## 2. Method

### 2.1. Materials and pretests

The materials used by MNC were adapted such that the relative clauses containing attractor were object-extracted, and the antecedent was the extracted object (see Table 1). There were three variants of relative clause – one where the antecedent was a verbal object (29 items), one where the antecedent was a prepositional object (79 items), and one where the antecedent was a prepositional object accompanied by a resumptive pronoun (12 items; see Table 1).

### 2.2. ERP experiment

#### 2.2.1. Participants

Thirty right-handed participants (7 males; mean age = 21 years) from the University of the Basque Country community gave written informed consent. All were native Spanish speakers with normal vision, and none had neurological or psychiatric disorders or had participated in the pretests.

#### 2.2.2. Procedure

The procedure was identical to MNC, involving presentation of 180 sentences (120 experimental, 60 fillers), with pre-critical through post-critical words presented for 300 ms on and 200 ms off, and comprehension questions following 25% of the sentences. Participants had good performance on the comprehension questions (mean percent correct across all conditions and fillers = 87%, SD = .08, participant range = 74–98%).

#### 2.2.3. Electroencephalogram recording

We used the same recording and preprocessing settings as MNC. We created an epoch of 1500 ms after the onset of the critical word, containing the critical word (*otra/a*) and the two words following it, which were the same across conditions. We considered the original 1200 ms epoch from MNC, but we extended that epoch by 200 ms because we observed the beginning of a large positive deflection that was cut off by the end of the MNC epoch.

#### 2.2.4. Statistical analysis

As in MNC, we performed a 2(Grammaticality: Correct, Incorrect) × 2(Attractor: Same, Different) repeated measures analysis of variance (ANOVA) using average amplitude across all electrodes in the 100–400, 400–1000, 1000–1200 ms time windows after critical word onset, and in the additional 1200–1400 ms time window. These windows are suitable to examine modulations of P600 effects induced by pronoun gender mismatch as well as of the Nref, which are both relatively extended ERP effects (see Nieuwland & Van Berkum, 2006, 2008b). Note that the 1000–1200 and 1200–1400 ms time windows correspond with P600 time windows on the post-critical word and that we label the time windows with reference to the critical word.

Following a significant interaction between scalp distribution and our factors, we performed the same analysis on anterior and posterior electrode groups, which also is consistent with the known distributions of ERP effects associated with referential ambiguity (anterior maximum of Nref effects) and grammatical gender violations (posterior maximum of P600 effects). We performed statistical tests using anterior (F3/4, F7/8, FC1/2, FC5/6, FP1/2, and Fz) and posterior (CP1/2, CP5/6, P3/4, P7/8, and Pz) channels, in addition to tests that involved all electrodes. Only statistically significant differences are reported. We also performed an ANOVA over all electrodes from 0–100 ms to check for potential baseline differences.

## 3. Results

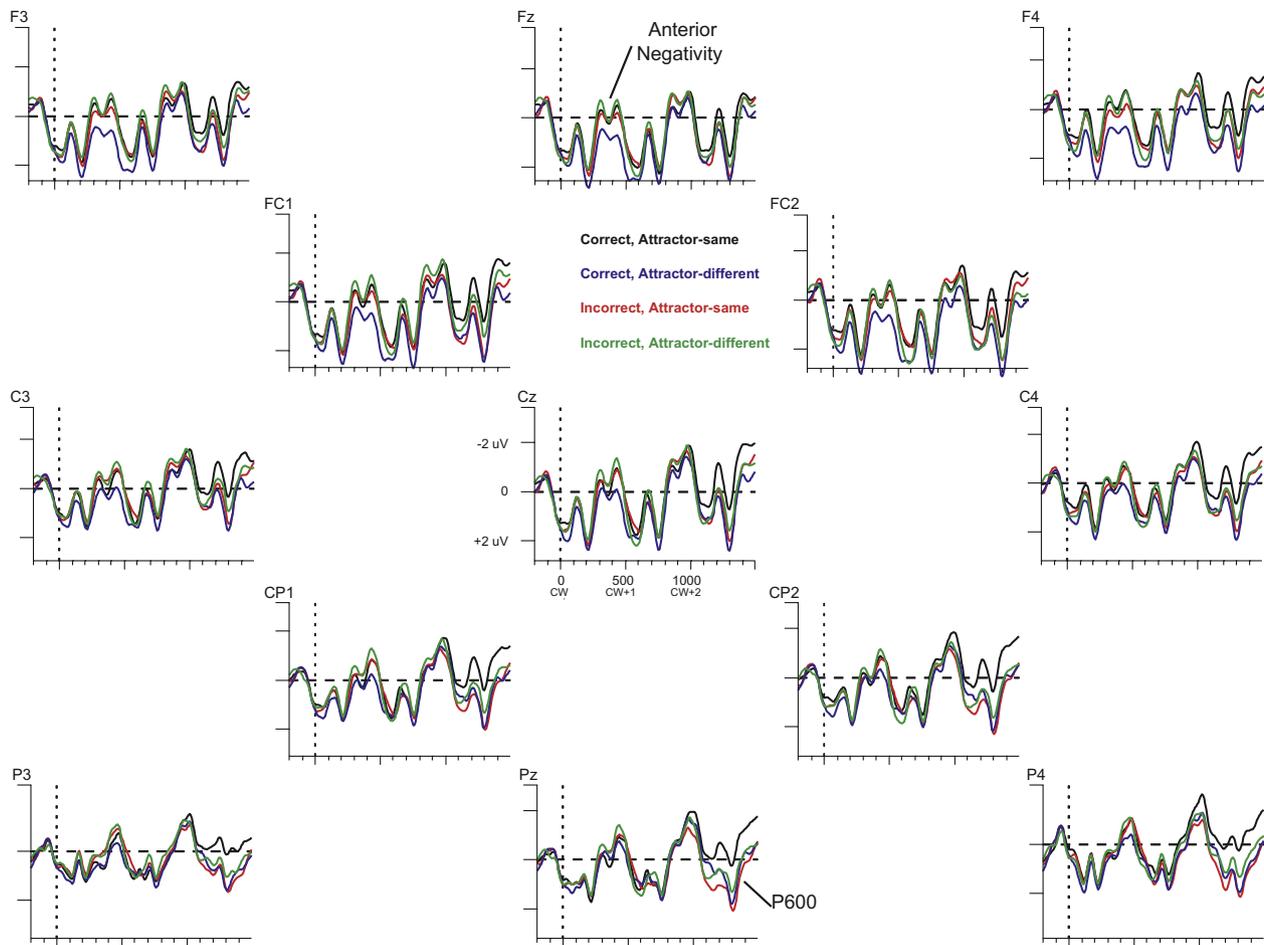
The ANOVA over all electrodes showed no significant effects ( $F_s < 2.1$ ) in any time windows (including 0–100 ms), except in the 1200–1400 window where there was a marginal Grammaticality × Attractor interaction ( $F(1,29) = 3.71, p = .06$ ). An analysis of scalp distribution via an ANOVA including anteriority/posteriority (AP) as a factor revealed a significant 3-way interaction between AP, Grammaticality, and Attractor in the 100–400 time window ( $F(1,29) = 10.12, p < .01$ ). As can be seen in Fig. 1, and in contrast to the MCN results, a relatively brief ERP modulation by the conditions occurred at anterior channels. Consistent with this pattern, a reliable Grammaticality by Attractor interaction occurred on anterior channels in the 100–400 ms window (see Fig. 1 and Table 2). Pairwise comparisons revealed that the Correct conditions were reliably different, with the Attractor-same condition being more negative (see Table 3). The Attractor-different conditions were also reliably different, with the Incorrect condition being more negative

**Table 1**  
Example sentences.

Condition	Example with antecedent as prepositional object without resumptive pronoun (79 items)											
1. Correct Attractor-same	Rafaela	perdió	el	<b>colgante</b>	que	junto	con	el	anillo	siempre	llevaba	y
	Rafaela	lost.3sg	the.m	necklace[m]	that	together	with	the.sg.m	ring [m]	always	wore.3SG	and
	Mónica	recuperó	<b>otro</b>	<b>que</b>	había	perdido	años	atrás.				
	Monica	recovered.3sg	another.sg.m	that	had.3sg	lost	years	back.				
	<i>'Rafaela lost the necklace that she always wore with the ring and Monica recovered another that she had lost years before.'</i>											
2. Correct Attractor-different	Rafaela	perdió	el	<b>colgante</b>	que	junto	con	la	sortija	siempre	llevaba	y
	Rafaela	lost.3sg	the.m	necklace[m]	that	together	with	the.sg.f	ring [f]	always	wore.3SG	and
	Mónica	recuperó	<b>otro</b>	<b>que</b>	había	perdido	años	atrás.				
	Monica	recovered.3sg	another.sg.m	that	had.3sg	lost	years	back.				
	<i>'Rafaela lost the necklace that she always wore with the ring and Monica recovered another that she had lost years before.'</i>											
3. <sup>~</sup> Incorrect Attractor-same	Rafaela	perdió	el	<b>colgante</b>	que	junto	con	el	anillo	siempre	llevaba	y
	Rafaela	lost.3sg	the.m	necklace[m]	that	together	with	the.sg.m	ring [m]	always	wore.3SG	and
	Mónica	recuperó	<b>otra</b>	<b>que</b>	había	perdido	años	atrás.				
	Monica	recovered.3sg	another.sg.f	that	had.3sg	lost	years	back.				
	<i>'Rafaela lost the necklace that she always wore with the ring and Monica recovered another that she had lost years before.'</i>											
4. <sup>~</sup> Incorrect Attractor-different	Rafaela	perdió	el	<b>colgante</b>	que	junto	con	la	sortija	siempre	llevaba	y
	Rafaela	lost.3sg	the.m	necklace[m]	that	together	with	the.sg.f	ring [f]	always	wore.3SG	and
	Mónica	recuperó	<b>otra</b>	<b>que</b>	había	perdido	años	atrás.				
	Monica	recovered.3sg	another.sg.f	that	had.3sg	lost	years	back.				
	<i>'Rafaela lost the necklace that she always wore with the ring and Monica recovered another that she had lost years before.'</i>											
	<u>Example with antecedent as verbal object (29 items)</u>											
1. Correct Attractor-same	Pedro	se	llevó	el	<b>coche</b>	que	el	todoterreno	golpeó	y		
	Pedro	refl.3sg	took.3sg	the.sg.m	car[m]	that	the.sg.m	SUV[m]	hit.3sg	and		
	Ander	compró	<b>otro</b>	solo	por	fastidiar.						
	Ander	bought.3sg	another.sg.m	only	to	annoy.inf.						
	<i>'Pedro took the car that the SUV hit and Ander bought another just to be annoying.'</i>											
2. Correct Attractor-different	Pedro	se	llevó	el	<b>coche</b>	que	la	moto	golpeó	y		
	Pedro	refl.3sg	took.3sg	the.sg.m	car[m]	that	the.sg.f	motorbike[f]	hit.3sg	and		
	Ander	compró	<b>otro</b>	solo	por	fastidiar.						
	Ander	bought.3sg	another.sg.m	only	to	annoy. inf.						
	<i>'Pedro took the car that the motorbike hit and Ander bought another just to be annoying.'</i>											
3. <sup>~</sup> Incorrect Attractor-same	Pedro	se	llevó	el	<b>coche</b>	que	el	todoterreno	golpeó	y		
	Pedro	refl.3sg	took.3sg	the.sg.m	car[m]	that	the.sg.m	SUV[m]	hit.3sg	and		
	Ander	compró	<b>otra</b>	solo	por	fastidiar.						
	Ander	bought.3sg	another.sg.f	only	to	annoy.inf.						
	<i>'Pedro took the car that the motorbike hit and Ander bought another just to be annoying.'</i>											
4. <sup>~</sup> Incorrect Attractor-different	Pedro	se	llevó	el	<b>coche</b>	que	la	moto	golpeó	y		
	Pedro	refl.3sg	took.3sg	the.sg.m	car[m]	that	the.sg.f	motorbike[f]	hit.3sg	and		
	Ander	compró	<b>otra</b>	solo	por	fastidiar.						
	Ander	bought.3sg	another.sg.f	<b>only</b>	to	annoy. inf.						
	<i>'Pedro took the car that the motorbike hit and Ander bought another just to be annoying.'</i>											
	<u>Example of a resumptive pronoun-containing sentence (antecedent as prepositional object: 12 items)</u>											
1. Correct Attractor-same	Marta	se	compró	la	<b>camiseta</b>	con	la	que	la	<u>falda</u>	iba	bien
	Marta	refl.3sg	bought.3sg	the.sg.f	t-shirt[f] <sub>i</sub>	with	it.sg.f <sub>i</sub>	that	the.sg.f	skirt[f]	went.3sg	well
	y	Miren	cogió	<b>otra</b> [...]	para	salir	de	fiesta.				
	and	Miren	took.3sg	another.sg.f	to	go_out.inf	of	party.				
	<i>'Marta bought the t-shirt that the skirt went well with and Miren took another to go out partying.'</i>											
2. Correct Attractor-different	Marta	se	compró	la	<b>camiseta</b>	con	la	que	el	<u>vestido</u>	iba	bien
	Marta	refl.3sg	bought.3sg	the.sg.f	t-shirt[f] <sub>i</sub>	with	it.sg.f <sub>i</sub>	that	the.sg.m	dress[f]	went.3sg	well
	y	Miren	cogió	<b>otra</b> [...]	para	salir	de	fiesta.				
	and	Miren	took.3sg	another.sg.f	to	go_out.inf	of	party.				
	<i>'Marta bought the t-shirt that the dress went well with and Miren took another to go out partying.'</i>											
3. <sup>~</sup> Incorrect Attractor-same	Marta	se	compró	la	<b>camiseta</b>	con	la	que	la	<u>falda</u>	iba	bien
	Marta	refl.3sg	bought.3sg	the.sg.f	t-shirt[f] <sub>i</sub>	with	it.sg.f <sub>i</sub>	that	the.sg.f	skirt[f]	went.3sg	well
	y	Miren	cogió	<b>otro</b> [...]	para	salir	de	fiesta.				
	and	Miren	took.3sg	another.sg.m	to	go_out.inf	of	party.				
	<i>'Marta bought the t-shirt that the dress went well with and Miren took another to go out partying.'</i>											
4. <sup>~</sup> Incorrect Attractor-different	Marta	se	compró	la	<b>camiseta</b>	con	la	que	el	<u>vestido</u>	iba	bien
	Marta	refl.3sg	bought.3sg	the.sg.f	t-shirt[f] <sub>i</sub>	with	it.sg.f <sub>i</sub>	that	the.sg.m	dress[f]	went.3sg	well
	y	Miren	cogió	<b>otro</b> [...]	para	salir	de	fiesta.				
	and	Miren	<b>took</b> .3sg	another.sg.m	to	go_out.inf	of	party.				
	<i>'Marta bought the t-shirt that the dress went well with and Miren took another to go out partying.'</i>											

(see Table 3). The 1000–1200 ms time window did not show a significant distribution interaction, and neither did the 1200–1400 ms time window, which had shown a marginal interaction on all electrodes. However, whereas P600 effects are usually maximal at posterior scalp location, they may fail to show statistically reliable differences between anterior and posterior locations due to the well-established broad

distribution of P600 effects. We therefore felt justified to proceed with a separate anterior and posterior electrode selection to obtain a more sensitive test of P600 modulations at posterior channels, thereby minimizing the chance that any obtained Grammaticality by Attractor interaction effect is driven by electrodes that do not show a substantial P600 modulation.



**Fig. 1.** Grand averages at all electrodes elicited by critical words: Correct attractor-same (*camiseta-falda-otra*; black); Correct attractor-different (*camiseta-vestido-otra*; blue); Incorrect attractor-same (*camiseta-falda-otro*; red); Incorrect attractor-different (*camiseta-vestido-otro*; green). Negativity is plotted upwards and that the waveforms are filtered (5 Hz high cut-off) for presentation purpose.

**Table 2**

F-values and MSE from ANOVA on average over anterior and posterior electrodes between 100–400, 400–1000, 1000–1200, and 1200–1400 ms ( $df = 1, 29$ ).

	Time (ms)			
	100–400	400–1000	1000–1200	1200–1400
<i>Anterior electrodes</i>				
Grammaticality				
F	1.56	.93	.30	.09
MSE	1.19	2.56	2.89	3.79
Attractor				
F	.83	.42	.39	1.37
MSE	1.18	2.25	2.55	3.79
Grammaticality $\times$ Attractor				
F	4.52*	1.63	1.90	2.22
MSE	1.38	2.2	2.73	4.45
<i>Posterior electrodes</i>				
Grammaticality				
F	2.61	.57	1.98	3.37 <sup>^</sup>
MSE	.73	1.66	2.12	1.83
Attractor				
F	.03	.06	.17	1.19
MSE	1.87	2.19	2.64	3.68
Grammaticality $\times$ Attractor				
F	.08	.36	.85	4.69*
MSE	.96	2.35	4.01	4.34

\*\* $p < .05$ .

\*  $p < .05$ .

<sup>^</sup>  $p < .08$ .

A later-occurring posterior positivity was observed in the Incorrect conditions compared to the Correct conditions, approximately 1100–1400 ms after CW onset, or between 600 and 900 ms after post-critical word onset, suggesting a P600 modulation by the post-critical word based on its timing and waveform morphology (e.g., Nieuwland & Van Berkum, 2006; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995). Statistical analysis revealed a reliable interaction between Grammaticality and Attractor on this posterior positivity in the 1200–1400 ms window (see Table 2), whereas the main effect of Grammaticality was marginally significant with Incorrect conditions more positive than Correct ones. Follow-up pairwise comparisons revealed a larger P600, compared to the Attractor-same Correct condition, for the Attractor-different Correct condition and for the Attractor-same Incorrect condition (see Table 4).

#### 4. Discussion

The current pattern of results shows interesting commonalities as well as differences with the results observed by MNC. Importantly, in terms of quantitatively different patterns across conditions, we replicated the main finding from MNC: the gender of a syntactically-unavailable attractor noun, intervening between ellipsis site and antecedent, impacted the electrophysiological response that accompanied ellipsis processing, even in fully

**Table 3**Pairwise comparisons on average voltage ( $\mu\text{V}$ ) over Anterior electrodes between 100 and 400 ms ( $df = 1, 29$ ).

	Incorrect	Correct	Incorrect – Correct (95% CI)
Attractor-same	.64	.44	.20 (–.42 to .84)
Attractor-different	.37	1.07	–.70** (–1.27 to –.14)
Attractor-different – Attractor-same (95% CI)	.27 (–.38 to .93)	–.63** (–1.17 to –.11)	

\* $p < .05$ .\*\* $p < .02$ .**Table 4**Pairwise comparisons on average voltage ( $\mu\text{V}$ ) over Posterior electrodes between 1200 and 1400 ms ( $df = 1, 29$ ).

	Incorrect	Correct	Incorrect – Correct (95% CI)
Attractor-same	.90	–.38	1.28** (.35 to 2.20)
Attractor-different	.45	.83	.37 (–.56 to 1.3)
Attractor-different – Attractor-same (95% CI)	.45 (–.53 to 1.41)	–1.21* (–2.35 to –.06)	

\*\* $p < .02$ .\* $p < .05$ .

grammatical circumstances. In contrast to our previous findings involving a sustained central-anterior negativity, however, the current effects were observed on an anterior negativity that occurred earlier and more briefly, and on a later-occurring posterior positivity that we take to be a P600 effect time-locked to the post-critical word. The observed patterns for each of these two ERP effects are discussed before turning to the more general implications of our findings for the notion of retrieval interference and for the neurobiology of language.

#### 4.1. Anterior negativity modulations on the critical word

Although there was no main effect of Grammaticality on the early negativity, there was a reliable difference between the Attractor-different conditions at anterior channels, with the Incorrect cue condition being more negative. This pattern is in contrast with the difference between Attractor-same conditions reported by MNC. Notably, the current stimuli feature a distinct syntactic role for the antecedent and attractor, where the antecedent was never a subject and could therefore be less prominent (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009). Crucially, however, we nevertheless observed a reliable difference between the correct conditions, as in MNC, albeit with the Attractor-same condition eliciting the main deviant, less negative ERP response. We speculate that the difference between the two studies in the pairwise effects stem from the distinct syntactic roles of attractor and antecedent, combined with the role of morphosyntactic gender information during retrieval. Although *otra* does not specify a particular syntactic role for its antecedent, those syntactic role features may still be represented along with the nouns in memory, contributing to their distinctiveness. Under this circumstance, we can observe similarity-based interference as a function of grammatical gender. We originally predicted this pattern in MNC: same-gender attractors in memory generate retrieval interference as compared to different-gender attractors, because gender morphology at the ellipsis site is no longer diagnostic to a unique antecedent. This prediction was not

borne out in MNC, where the antecedent was the subject of the relative clause and where mismatching gender on the attractor noun disrupted retrieval. However, it was borne out in the anterior negativity effect in the current study, signaling that where syntactic role information is arguably distinctive between antecedent and attractor, gender cue-driven similarity-based interference appears to emerge.

Interestingly, the anterior negativity observed on the critical word was earlier in onset and more anterior than the negativity observed by MNC. These features are, however, more in line with those of the Nref effect observed for referential processing difficulty (see Nieuwland & Van Berkum, 2008a; Van Berkum et al., 1999). We cannot adequately explain the apparent difference between the studies in the effect distribution, nor in effect timing. However, as to why this effect was not of a sustained nature in the current study, we did observe positive deflections in a subset of participants 400–1000 ms time window (for Incorrect versus Correct), and it is possible that these positive ERP responses capped the sustained nature of the observed negativity. We now refer the reader to the Appendix A, in which we examined the contribution of individual differences by splitting the participants into two groups based on the main effects of grammaticality, in absence of an *a priori* reason to group participants (see also Nieuwland & Van Berkum, 2008b). This analysis revealed that participants who showed a negativity to incorrect conditions compared to correct conditions in fact showed a sustained negativity difference between the two grammatical conditions. While the currently obtained negativity contrasts with the MNC results in that it was larger for the attractor-same conditions, the P600 component that occurred at the post-critical word did replicate the MNC results in terms of condition patterns, which we turn to now.

#### 4.2. P600 modulations on the post-critical word

The observed P600 modulation was typical in terms of scalp distribution and waveform morphology (e.g., Osterhout &

Holcomb, 1992; Osterhout & Mobley, 1995). This confirms the prediction from MNC and from the current study, based on the extant ERP literature on syntactic processing (e.g., Barber & Carreiras, 2005; Osterhout & Mobley, 1995; Silva-Pereyra et al., 2012), that a larger P600 would be elicited by the incorrect conditions compared to the correct conditions, reflecting reprocessing or repair of the incorrect ellipsis. Interestingly, although the current P600 modulation does not literally replicate the observed ERP effects in MNC, the differential pattern between conditions was the same as in MNC: a larger P600 for the Incorrect Attractor-same condition compared to the Correct Attractor-same condition, and a larger P600 for the Correct Attractor-different condition compared to the Correct Attractor-same condition. These results thus align with those reported by MNC that the gender of the attractor influenced the processing of the ellipsis, even in fully grammatical sentences.

A further unexpected finding was the late onset of the P600 modulations. In fact, we take the latency of the current P600 modulation to reflect its elicitation by the post-critical word (with the peak of the effect approximately 700–900 ms post-critical word onset), rather than by the critical word. Although this assumption does not directly bear upon the P600 differences between conditions per se, we can speculate why this effect might be time-locked to the post-critical word. It is possible that upon encountering *otra/a*, rather than immediately attempt to resolve the ellipsis, participants generated an expectation for an upcoming noun phrase (e.g., *otra camiseta*), despite the fact that none of the sentences in the experiment contained this type of utterance. Only when participants encounter the next word their expectation is falsified (*para* rather than a new noun phrase), which is essentially a syntactic preference violation ('garden-path'; Osterhout, Holcomb, & Swinney, 1994). However, this explanation does not fit easily with the early negativity modulations that occurred before presentation of the post-critical word, nor with the MNC results that also showed retrieval processes being elicited before the post-critical word.

#### 4.3. Composite retrieval cues during sentence comprehension

Our results suggest that the degree to which retrieval cues match or 'pick out' a unique antecedent in memory drives neural responses during dependency resolution, and that this degree of match, or *diagnosticity*, appears to be calculated over composite sources of information. Whether a set of retrieval cues, or a composite retrieval cue, is diagnostic to a target in memory depends on the retrieval cue's resonance with the target relative to its resonance other items in memory (Gillund & Shiffrin, 1984; Nairne, 2002; see Van Dyke & McElree, 2011). Although in principle including morphosyntactic gender in a composite retrieval cue seems helpful, it does not automatically increase diagnosticity. In our materials, *otra* is in fact always diagnostic to its antecedent *camiseta* based on structural position alone, one could argue that morphosyntactic gender information therefore would not contribute to higher diagnosticity. Nonetheless, the current and previous findings suggest that morphosyntactic gender information from syntactically unavailable nouns impacts retrieval. Instead then, these two pieces of evidence seem to suggest that morphosyntactic gender information, typically a robust source of information or distinctiveness of nouns in Spanish, cannot be excluded from the information interacting with memory at retrieval, even if it is irrelevant. Our results, and those reported by MNC, seem to suggest that different types of syntactic information, syntactic role and morphosyntactic gender, play a role in eliciting the antecedent from memory.

#### 4.4. Agreement attraction in grammatically correct sentences

Our results replicate the MNC findings with regard to the intrusion of grammatically unavailable nouns on the resolution of referring expressions in grammatically correct sentences. Moreover, they extend the pattern of interference effects to a different syntactic context – tentatively suggesting that distinctive syntactic role information changes the ERP morphology of the interference effects. Our results thus contrast with previous ERP findings on subject–verb agreement attraction effects during comprehension, where smaller P600 effects have been reported for subject–verb violations in the context of a locally-matching attractor (e.g., "The key to the cabinets are ...", Tanner et al., 2012; although see Kaan, 2002). Our findings thus also contrast with findings from the behavioral literature, where agreement attraction has been reported to occur for subject–verb agreement violations but not for reflexive pronoun gender-agreement agreement violation (e.g., Dillon et al., 2013; Sturt, 2003; Wagers et al., 2009; see Phillips, Wagers, & Lau, 2010, for review), with the notable exception of Cunnings and Felser (2013) who reported agreement attraction effects in low-span individuals for grammatically-correct reflexive pronouns. While it is interesting to speculate why we found effects that are conceptually potentially very different from the previous ERP and behavioral studies on agreement attraction (effectively, an illusion of *ungrammaticality* as opposed to the previously reported illusion of *grammaticality*), we cannot overstate the differences between our paradigm and previous work. Beyond differing in language (Spanish versus English), agreement type (grammatical gender agreement versus number agreement) and task (manipulation-irrelevant comprehension questions versus grammaticality judgment), while the agreement attraction literature predominantly involves short-distance subject–verb number agreement (as inspired by agreement production errors, see Bock & Miller, 1991), our paradigm involves a referential manipulation (i.e., introducing a new referent into the discourse model) with a large distance of about 7–10 words between the agreeing determiners and antecedents. Furthermore, in contrast with previous ERP and behavioral work on agreement attraction, our study did not require participants to evaluate grammaticality explicitly via a task, which clearly might have implications for the type of processing strategies that participants engage in during an experiment. In sum, although the interference effects we report share core theoretical notions with the above recently reported studies on number attraction, the fact that they were found via a comparison of grammatical sentences sets them apart. Further research is needed to determine how differences in experimental paradigm and linguistic cues in different languages contribute to the morphology of interference effects.

Additionally, the fact that the Correct Attractor-same condition was more negative than the Correct Attractor-different condition suggests that when the attractor had the same gender as the antecedent, this similarity was disruptive to processing. This pattern is consistent with similarity-based interference (e.g., Gordon, Hendrick, Johnson, & Lee, 2006; McElree, 2006; McElree et al., 2003; Van Dyke, 2007; Van Dyke & McElree, 2011), as well as agreement attraction from syntactically unavailable constituents (e.g., Wagers et al., 2009). However, it is in contrast with previous findings from MNC. The interaction pattern on the post-critical word in the current study, however, does replicate the interaction effect found in MNC, where the Correct Attractor-different was deviant compared to the Correct Attractor-same condition. Tentatively, similarity of the antecedent and attractor (*otra* eliciting *camiseta* in the context of *falda*) and diagnosticity of the cue to the antecedent relative to the attractor (*otra* eliciting *camiseta* in

the context of *vestido*) both can shape processing of dependencies such as ellipsis, with as yet unspecified contextual variables tipping the balance between one or the other. Though the diagnosticity of a retrieval cue to its target is can be a function of similarity, it may also take into account the cue's resonance or link to other items in memory (Martin et al., 2012; Nairne, 2002). More research is needed to map those contextual factors and to ascertain if these variables come into play on different timescales during processing.

#### 4.5. Conclusions

We found further ERP evidence that retrieval during sentence comprehension can be subject to interference from intervening representations, even if they are structurally unavailable for ellipsis and the sentence is grammatically correct. Importantly, our results are consistent with cue-dependent direct-access retrieval of content-addressable representations underlying online language comprehension, and emphasize further incorporation of such computational principles and architectures into theories of language processing.

#### Acknowledgments

We thank Eri Takahashi, Saioa Larraza, and Brendan Costello. AEM was supported by a Juan de la Cierva fellowship from the Spanish Ministry of Science and Innovation (MICINN) and a ESRC Future Research Leaders fellowship. This work was supported by the Economic and Social Research Council [Grant number RA2553], to AEM. MSN was supported by a Ramon y Cajal fellowship and a Plan Nacional research grant from MICINN, and MC was supported by CONSOLIDER-INGENIO2010CSD2008-00048 from MICINN.

#### Appendix A. Individual differences on the critical word (CW)

Substantial inter-subject variability was present in the brain response to the critical word, with some individuals showing P600 modulations at the CW rather than the negativity observed for the group as a whole. In a group-split analysis, we explored whether the short duration of the negativity seen reported in the main text, compared to the sustained negativity found in Martin, Nieuwland, and Carreiras (2012; henceforth MNC), may have been due to an overlapping P600 response from some participants.

Our approach is similar to that taken by Nieuwland and Van Berkum (2008a,b), who also did not have an *a priori* way to divide participants into groups (see also Osterhout, 1997): splitting the subjects on a variable that is orthogonal to the variable of interest, which allows one to investigate if the two pattern together. We computed the main difference between Correct cue and Incorrect cue conditions in the 400–1000 ms time window (MNC; Nieuwland and Van Berkum, 2008a,b), and created two groups based on this difference: one whose mean difference was negative (henceforth the Nref group,  $N = 18$ ) and one whose mean difference was positive (P600 group,  $N = 12$ ). This split is likely an oversimplification of the data, but a quantitative splitting criterion was more appealing than grouping individual averages based on visual inspection. Notably, this group-split renders the main effects trivial, but because it is orthogonal to the difference between the two grammatical conditions it can thus be used to assess whether group differences in the overall response pattern affected the duration of the negativity between the grammatical conditions on the CW.

In our overall analysis, it was the reliable difference between the two grammatical conditions that drove the interaction pat-

terns across all subjects (see also MNC). Given this finding and the theoretical importance of a difference between grammatical conditions, we conducted further statistical tests only on these two conditions between 400–1000 and 1200–1400 ms as a function of group. We ask: Does the Nref group show a sustained differential effect of attractor in the grammatical conditions (as seen across all participants)? Does the differential effect of attractor in grammatical conditions on the later P600 occur whether or not participants show an earlier negativity for the main effect?

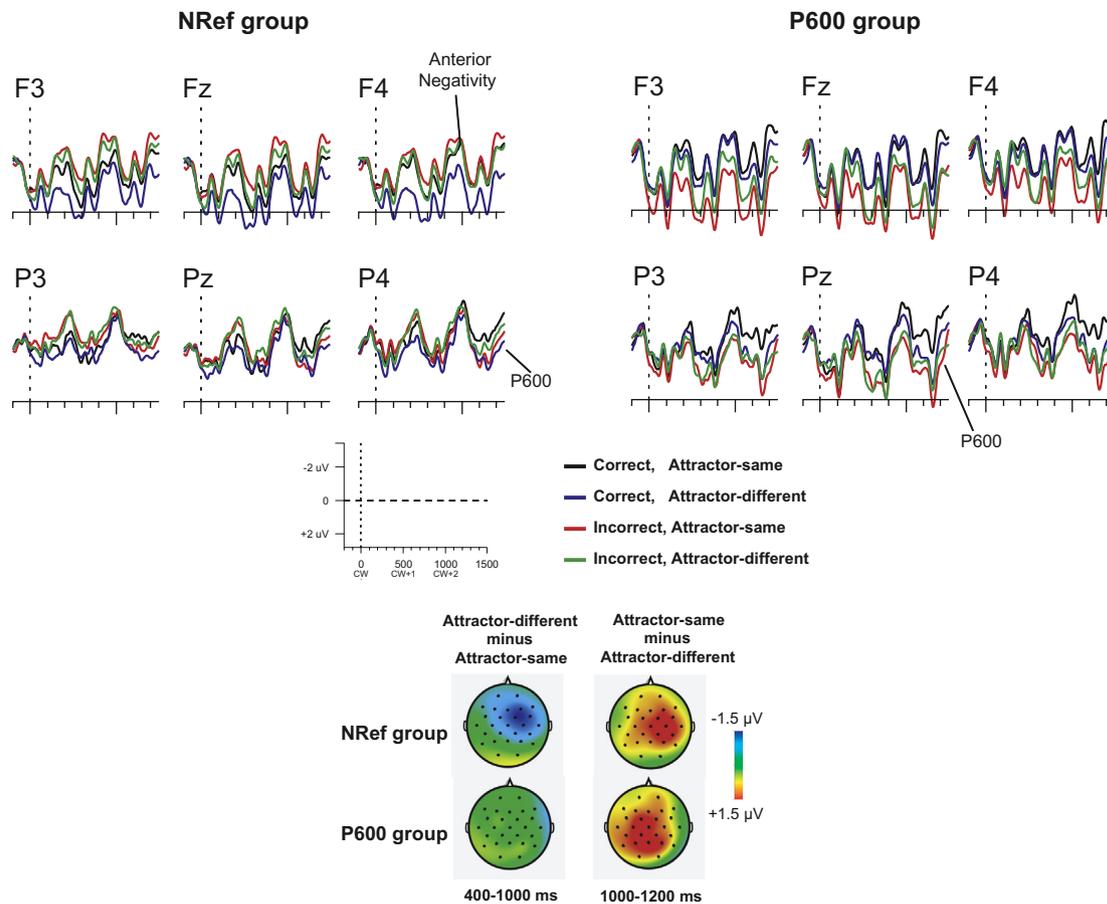
The graphs in Appendix B show the grand average ERPs for each group and scalp distribution of the difference between the two correct-cue conditions in both time windows. We conducted a repeated measures Distribution (Anterior, Posterior)  $\times$  Attractor (Same, Different) ANOVA on the Correct cue conditions with Group (Nref, P600) as a between-subjects factor on 400–1000 ms. We found a reliable Distribution  $\times$  Attractor  $\times$  Group interaction ( $F(1,28) = 4.76$ ,  $p < .05$ ), reflecting the anterior negativity seen in the Nref group but not in the P600 group. We followed up this interaction with a Distribution  $\times$  Attractor repeated measures ANOVA within each group. In the Nref group, there was a reliable Distribution  $\times$  Attractor interaction ( $F(1,17) = 7.15$ ,  $p < .05$ ). Following up on this interaction, we analyzed anterior channels of the Nref group and found a reliable difference between the Grammatical conditions, such that the Attractor-Same condition was more negative than the Attractor-Different condition,  $F(1,17) = 7.55$ ,  $p < .01$ . There were no effects on posterior channels for the Nref group. The P600 group did not show any reliable effects. These effects suggest that the Nref group is showing a reliable difference between grammatical conditions between 400 and 1000 ms, the same time window as in MNC, but that, in combination with the P600 group, the duration of the effect is shortened on the grand average over all participants.

We then performed a repeated measures Distribution (Anterior, Posterior)  $\times$  Attractor (Same, Different)  $\times$  Group (Nref, P600) on 1200–1400 ms post-CW. No reliable by group interaction effects were found (all  $F_s < .5$ ). Although absence of a Group effect must be interpreted with caution (of note, this pattern remained unchanged when using a posterior channel selection where P600 modulations are most pronounced, which would be more sensitive to pick up group differences), we take it to reflect that the two groups show a similar pattern of activity between 1200 and 1400, with a larger P600 modulation for the Correct Attractor-Different condition than for the Correct Attractor-Same condition.

In sum, our group-split analysis revealed two novel findings: (1) The Correct Attractor-Same condition indeed elicited a sustained frontal negativity compared to the Correct Attractor-Different condition, but this effect was only present in participants who showed a negativity on the main effect of Grammaticality for Incorrect cue versus Correct cue conditions. This finding deviated from MNC in that now the Correct Attractor-Same condition patterns with the Incorrect cue conditions. (2) The P600 modulation as seen on the post-critical word appeared unaffected by these earlier individual differences, as both groups showed a larger P600 for the Correct Attractor-Different condition than for the Correct Attractor-Same condition. The latter finding seems consistent with the MNC results in that the Correct Attractor-Different condition patterns with the Incorrect cue conditions.

#### Appendix B

See Fig. B1.



**Fig. B1.** Upper graphs: Grand average ERPs for the NRef group ( $N = 18$ ) and P600 group ( $N = 12$ ). Group split was based on the average voltage between 400 and 1000 ms for incorrect conditions minus correct conditions across all channels (negative-Nref group, positive-P600 group). Lower graphs: Scalp distributions for the differences between the two grammatically correct conditions. In the 400–1000 ms time window, the NRef group shows a larger frontal negativity for attractor-different versus attractor-same. In the 1000–1200 ms time window, both groups show a larger P600 component for attractor-same compared to attractor-different.

## References

- Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and memory. Memory* (Vol. 10, pp. 237–313). San Diego: Academic Press.
- Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28, 748–769.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17(1), 137–153.
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23(1), 45–93.
- Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113, 787–821.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*, 3, 19–58.
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98, 159–168.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, 28(1–2), 188–219.
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*.
- Eguren, L. (2010). Contrastive focus and nominal ellipsis in Spanish. *Lingua*, 120, 435–457.
- Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, 39(2–3), 225–242.
- Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 32(6), 1304.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1–65.
- Hanulíková, A., Van Alphen, P. M., Van Goch, M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, 24(4), 878–887.
- Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 273–316). Oxford, UK: Oxford University Press.
- Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject–verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, 31(2), 165–193.
- Kaan, E., Dallas, A. C., & Barkley, C. M. (2007). Processing bare quantifiers in discourse. *Brain Research*, 1146, 199–209.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98–110.
- King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376–395.
- Martin, A. E., & McElree, B. (2008). A content-addressable pointer underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, 58, 879–906.
- Martin, A. E., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35, 1231–1239.
- Martin, A. E., & McElree, B. (2011). Direct-access retrieval during sentence comprehension: Evidence from sluicing. *Journal of Memory and Language*, 64, 327–343.
- Martin, A. E., Nieuwland, M. S., & Carreiras, M. (2012). Event-related brain potentials index cue-based retrieval interference during sentence comprehension. *NeuroImage*, 59, 1859–1869.
- McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46). San Diego: Academic Press.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48, 67–91.
- Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, 10, 389–395.

- Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, 36(4), 569–587.
- Nieuwland, M. S., & Van Berkum, J. J. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, 1118(1), 155–167.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2008a). The neurocognition of referential ambiguity in language comprehension. *Language and Linguistics Compass*, 2(4), 603–630.
- Nieuwland, M. S., & Van Berkum, J. J. (2008b). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, 106(2), 119–131.
- Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, 59(3), 494–522.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 786.
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41(3), 427–456.
- Phillips, C., Wagers, M., & Lau, E. (2010). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (Ed.), *Experiments at the interfaces. Syntax & Semantics* (Vol. 37). Bingley, UK: Emerald Publications.
- Silva-Pereyra, J., Gutierrez-Sigut, E., & Carreiras, M. (2012). An ERP study of coreference in Spanish: Semantic and grammatical gender cues. *Psychophysiology*, 49, 1401–1411.
- Suñer, M. (1998). Resumptive restrictive relatives: A crosslinguistic perspective. *Language*, 74, 335–364.
- Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, 60(2), 308–327.
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542–562.
- Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *Proceedings of the 36th Boston university conference on language development* (pp. 594–606). Somerville, MA: Cascadia.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. M. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41, 147–182.
- Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, 1146, 158–171.
- Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33, 407–430.
- Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay of attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, 49, 285–413.
- Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, 65, 247–263.
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237.