Journal of Neurolinguistics 68 (2023) 101154

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

The processing of *Which* interrogative sentences: A behavioral and ERP study☆

Mirko Grimaldi [a,*,1], Marica De Vincenzi [b,c], Paolo Lorusso [d], Francesco Di Russo [e,f], Rosalia Di Matteo [b], Luigi Rizzi [g], Maria Teresa Guasti [h]

[a] *Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL) – University of Salento, Lecce, Italy*
[b] *University G. d'Annunzio of Chieti-Pescara, Italy*
[c] *Marica De Vincenzi' Foundation ONLUS, Italy*
[d] *University of Udine, Italy*
[e] *University of Roma "Foro Italico", Italy*
[f] *IRCCS Fondazione Santa Lucia, Roma, Italy*
[g] *Collège de France, France*
[h] *University of Milano-Bicocca, Italy*

ARTICLE INFO

Keywords:
Wh-questions in Italian
Which-noun interrogatives
Syntactic parsing
Subject-object asymmetries
P600
Source analysis

ABSTRACT

This study investigates the parsing of Italian *Wh*-questions of the *Which-N* type. The extraction site could be either the subject or the object noun phrase. The verb following the *Which*-noun was either a singular or a plural form, immediately disambiguating the *Which-N* argument role through verb agreement. Reading time on the verb and on the post-verbal noun phrase were significantly shorter for the subject *Wh*-question than for the object *Wh*-question. Multi-channel ERP data showed increased P600 amplitudes for the object questions in response to the critical word on the left temporal lobe in the superior temporal gyrus. These findings are in line with the Minimal Chain Principle (De Vincenzi, 1991a) and provide further evidence for the hypothesis that the amplitude and duration of the P600 involve multi-dimensional processes controlling operations such as prediction, retrieval, revising, and structure-building operations needed for assembly (and disassembly) of syntactic relations.

1. Introduction

1.1. Differences in Wh-extraction: Theoretical perspectives

A crucial feature of natural languages is the possibility to combine words into hierarchically organized sentence structures and generate dependencies between non-adjacent words and structural chunks. A particular kind of long-distance dependency is

☆ This research was designed by Marica De Vincenzi in the last months of her life. After completing the experimentation and analysis, the coauthors felt that the results are novel and relevant enough to be made available to the international community of psycholinguists, after a necessary but minimal updating of theoretical background and methodology. We intend to dedicate this publication to Marica's memory.

* Corresponding author.
E-mail address: mirko.grimaldi@unisalento.it (M. Grimaldi).

[1] Mirko Grimaldi and Marica De Vincenzi are equal contributors to this work and designated as co-first authors. Corresponding author: Mirko Grimaldi, Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL) – University of Salento (73100, Lecce, Italy). mirko.grimaldi@unisalento.it.

https://doi.org/10.1016/j.jneuroling.2023.101154
Received 13 January 2023; Received in revised form 2 June 2023; Accepted 5 June 2023
Available online 4 July 2023
0911-6044/

represented by *Wh-* questions such as *Which mayor do the politicians attack?* In these cases, the direct object of the verb is moved to the front of the sentence so that the grammatical function and thematic role of the *Wh-* element cannot be fully determined until parsing operations establish a syntactic dependency between 'Which mayor' and the verb 'attack': this leads to the identification of 'Which mayor' as the syntactic object of the verb to which the verb assigns a thematic role (*patient* in this case). Italian raises additional issues in this connection, as it allows postverbal subjects. Whereas in English, subject and object questions are disambiguated by word order (*Which mayor attacks the politicians?* Vs *Which mayor do the politicians attack?*), in Italian, the sequence *Which* N – V – DP is fully ambiguous between a subject and an object question interpretation (*Quale sindaco ha attaccato l'uomo politico?* can mean both "Which mayor attacked the politician?" and "Which mayor did the politician attack?"). From a neurocognitive point of view, the challenge is to find out what processing steps the speakers go through to understand these sentences. In particular, the following questions arise (1) How quickly does brain processing (i.e., the parser): find a representational solution for these grammatical roles? (2) What formal and interpretive properties does the parser use? (3) Is there an effect of grammatical role? (4) Does the parser treat *Which*-questions in a special way, given their structural and interpretive properties (their lexically-restricted, discourse-linked character, see below)? In the present study on the parsing of subject/object *Wh-* elements the two nominal expressions selected by the transitive verbs mismatch in number, so that the ambiguous sequence is disambiguated by the verbal morphology, which always agrees with the subject. So, *Quale sindaco ha attaccato gli uomini politici* ("Which mayor$_{SING}$ attacked$_{SING}$ the politicians") can only be interpreted as a subject question since the *Wh-* element agrees with the verb, whereas *Quale sindaco hanno attaccato gli uomini politici?* ("Which mayor$_{SING}$ attacked$_{PLUR}$ the politicians$_{PLUR)}$?" = Which mayor did the politicians attack?) can only be interpreted as an object question since the postverbal DP *the politicians* agrees with the verb, hence it must be interpreted as the subject.

Several reading time experiments conducted in Dutch, English, Italian, German have already shown that a *Wh-* element tries in general to be immediately posited as filling the first argument position in a sentence (Frazier & Flores D'Arcais, 1989; Frazier & Clifton, 1989; Tanenhaus et al., 1989; Farke & Felix, 1994, Schlesewsky et al., 1999, Friederici, 2002, Penolazzi et al., 2005). The data on Italian (De Vincenzi, 1991a,b, Pagliarini et al., 2013) also show that there are differences between *Who* and *Which-N* questions. However, we will not address this aspect here, but rather focus only on *Which*-N questions. In the present study, we will address the issue of how the disambiguation of ambiguous object/subject *Which-Ns* through number morphology involves differences in both behavioral and electrophysiological processes.

1.2. Which-N extraction Subject-Object asymmetries: a neurophysiological account

The aim of the study is to investigate the parsing of subject/object *Wh-* questions disambiguated through the number marking on the verb that immediately follows the *Which-N*. When the two DPs are found in the order *Which*-N + V + DP and do not differ in number, as in (1), the interpretation for the *Which-N* is ambiguous between a subject (a) or an object (b) reading, since Italian allows postverbal subjects such as *the horse* in (1b).

(1)	Quale mucca insegue il cavallo?	
	Which cow follows the horse?	
	a. Which cow$_i$ t$_i$ follows the horse	Subject Reading (The cow follows the horse)
	b. Which cow$_k$ does the horse follow t$_k$	Object Reading (The horse follows the cow)

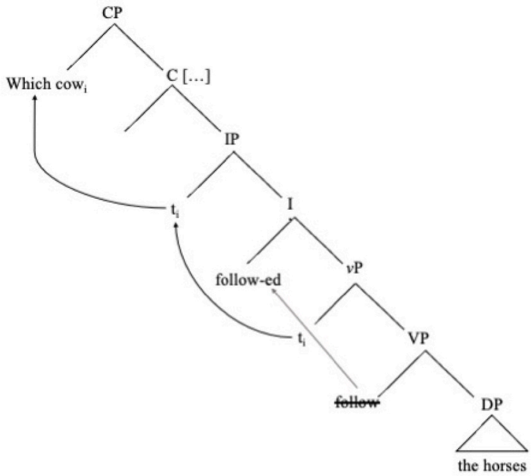
The most relevant grammatical issues involved in parsing sentences like (1) are: the locus of extraction of the *Which-N*, the D-linked status of the *Wh*-element, and the length of the DP movement. As for the *loci of extraction,* as we already mentioned, in Italian sentences like (1) they are ambiguous and allow **for** two readings: a subject (a) and an object (b) reading.

However, when the extracted DPs and the postverbal DPs mismatch in number, the sentences are disambiguated by the inflected verbs which agree with the subject. So, when the verb (e.g., *inseguì*) agrees in number with *Which-N*, the interrogative element is unambiguously interpreted as the subject (2) and is extracted from the subject position. On the contrary, when the verb (e.g., *inseguirono*) agrees in number with the postverbal DP, the postverbal element is unambiguously interpreted as the subject while the *Which-N* is interpreted as the object and is extracted from the complement position in the VP (2–3).[2]

(2)	Subject *Which-N*			
	Quale	*mucca*	*inseguì*	i cavalli?
	Which cow	*followed*$_{3\ sing}$	the horses?	
	'Which cow followed the horses?'			

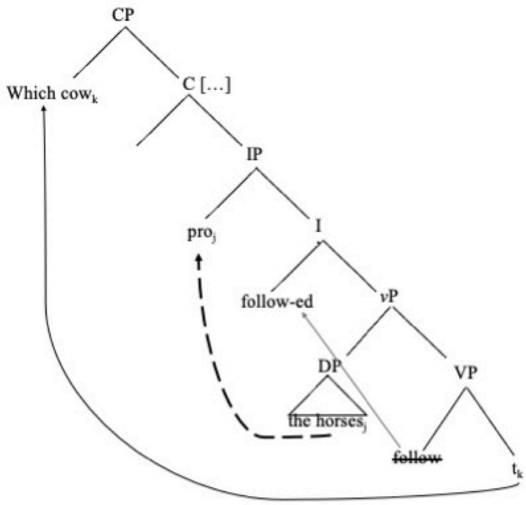
(3)

[2] In the following tree structures (3) and (5) we use English morphemes for ease of reading to express Italian structures. It should be noticed that in Italian the lexical verb moves from the VP to get associated with the inflection (a movement expressed by a straight line in (3) and (5)), whereas in English it does not (Pollock, 1989). The dotted line in (5) expresses the relation between the expletive null subject *pro* and the postverbal subject, a relation which is assimilated to a chain in Chomsky (1981) and Burzio (1986).



(4)	Object *Which-N* Quale mucca *inseguirono* Which cow *followed*$_{3pl}$	 *i cavalli*? *the horses*?

(5)



Given the fact that a single multimembered chain is involved in (3), whereas two multimembered chains must be postulated in (5) (expressed by the solid and dotted lines) this difference can be expected to have an impact on processing, in ways observable in behavioral and neurolinguistic data, as predicted by the Minimal Chain Principle (De Vincenzi, 1991a): henceforth MCP. The MCP is a processing principle based on the idea that chains are costly structures to maintain in short term memory. The parser should: (i) avoid postulating an unnecessary chain, but (ii) it should not delay postulating a chain once an element is identified as moved from somewhere else. The MCP predicts a preference for parsing *Wh*-N-V-DP structures as subject questions which require a single (multimembered) chain for the subject, as in structure (3). So, *quale mucca* in (2) and (6a) will be correctly taken as the subject of the verb *follows*. On the contrary, in (4) and (6b) the subsequent input word, the plural form of *follow*, signals that *quale mucca* cannot be the subject, requiring reanalysis and ultimately leading to a representation like (5). Since reanalysis results in longer reading times and in a specialized ERP component, we expect a longer reading time both on the plural verb and on the post-verbal noun. The postverbal position could actually be an extraction site for the object when no overt subject is present(6b) (recall that Italian allows subjectless sentences).

(6)	a	*Wh*-subject Quale mucca__insegue i cavalli nella savana? Which cow follows the horses in the wild?

(*continued on next page*)

(*continued*)

		'Which cow follows the horses in the wild?'
	b	*Wh*-object Quale mucca inseguono i cavalli ___nella savana? Which cow follow the horses in the wild? 'Which cow are the horses following in the wild?'

In the literature on language acquisition and language pathologies, another crucial grammatical factor has been brought to bear on the complexity differences between (6a) and (6b): the postverbal subject intervenes in the computation of the object chain in (6b), whereas there is no intervention in the case of (6) a. The *Which-N* is lexically restricted (contains a noun, as in *Which cow*, for example); so, when the intervening element is also lexically restricted (the horses), the intervener is of the same structural type as the target of the dependency. This produces a complexity effect that cannot be overcome by certain populations: young typically developing children (Friedmann et al., 2009; Guasti, 2017), SLI children (Friedmann & Novogrodsky, 2011) and agrammatic aphasics (Grillo, 2005, 2008). This line of research, inspired by the Relativized Minimality tradition in formal linguistics (Rizzi, 1990, 2004) makes distinct predictions with respect to the MCP in cases in which the *Wh*-element and the intervener are different in internal structure (e.g., a bare *Wh*-object and a lexically restricted intervening subject). As the data tested in the research reported in the present paper do not include this case, we continue to phrase our analysis in terms of De Vincenzi's MCP, leaving for future work a fuller discussion of the respective predictions and the compatibility of the MCP with an approach based on intervention.

1.3. Subject object asymmetries: neurophysiological perspective

In order to address the Subject-Object asymmetries with *Which-Ns*, a reading time experiment will be replicated with an Event-related Potential (ERP) study. Using high-density electroencephalography, topographical mapping, and source analysis techniques, we aim to investigate the ERP components involved in the object extraction of the *Which*-N. ERP research has shown specific neural activity related to the progressive integration of word meaning into the development of the sentence. This neural activity, the N400 effect, is a negative component, generally widespread all over the scalp but more intense in the posterior areas that peaks at about 400 ms after stimulus onset (for a review, see Kutas & Van Petten, 1994; Osterhout & Holcomb, 1995; Kutas & Federmeier, 2011). With respect to syntactic processing, two ERP components have been identified: an early anterior negativity and a late centro-parietal positivity. Syntactic anomalies elicit a negative polarity between about 200 and 500 ms after the onset of the critical word: these negativities are usually largest over left anterior sites. An early LAN (also called ELAN) around 250 ms has been found in response to phrase structure violations (Friederici et al., 1993; Friederici & Weissenborn, 2007; Neville et al., 1991). A LAN between 300 and 500 ms has been reported in relation to subject-verb number agreement violations (e.g., De Vincenzi et al., 2003) or in general with agreement violations (see Molinaro et al., 2011 for a review), and verb sub-categorization information (Osterhout & Holcomb, 1992; Rösler et al., 1993; Holcomb, 1993; Friederici & Weissenborn, 2007).

The second component reported in relation to syntactic processes is a large-amplitude positivity, with a stronger activity on the centro-parietal areas generally labeled P600 that starts at 500–600 ms after presentation of the syntactically critical word and persists 500–600 ms longer. It is elicited by anomalies involving phrase structure, verb tense, subject-verb number agreement, case violations, and also by constituent movement and garden-path sentences (cf. De Vincenzi et al., 2003 and Gouvea et al., 2010 for a review on all the types of P600 including the one evoked by *Wh*-dependencies). These data suggest that this ERP component is sensitive to different syntactic sub-processes, such as structural repair, and reanalysis. Indeed, recent findings have shown that the P600 is sensitive to semantic information. Specifically, it seems that semantic expectancy, thematic role and semantic reversal anomalies, as well as semantic incongruences on the sentence and discourse levels, may modulate this component. Moreover, P600 effects were reliably noted with both syntactically and semantically well-formed sentences where the pragmatic/conceptual content was manipulated, such as for figures of speech (hyperboles, metaphors, etc.), jokes, in response to sentence verification tasks, and discourse comprehension (cf. Friederici, 2002, 2017; Tanner et al., 2017; Regel et al., 2014 for an analysis of P600 and neural oscillation).

To summarize, ERP data suggest that a significant piece of the syntactic processes begin early with the E/LAN (at 300 ms), the stage when computational operations are carried out at the lexical level. Then, syntactic, semantic, and prosodic information is integrated, so that the P600 amplitude may reflect reanalysis and reparation processes necessary to globally compute the relations (structure building) between words within a sentence reconciling multiple (mismatching) representations in memory (cf. Friederici & Kotz, 2003 for a review, & Tanner, Gray, Hell, 2017, Tanner, 2019 for a critical approach). On the other hand, the latency of the P600 may reflect the time needed to retrieve the elements that participate in a structural relation (Gouvea et al., 2010).

The ERP studies that addressed a constituent movement relationship are limited, and, to our knowledge no studies have been performed exclusively on *Which*-N subject/object asymmetries. We will review different studies on the ERP response to *Wh*- fronting.

Seminal research by Kaan et al. (2000) compared sentences containing fronted direct object *Wh*-phrases (7a-b) with sentences in which the direct object appears in its canonical position (7c), and showed that a P600 response was elicited at the verb position (underlined) in the *Wh*–fronting conditions, relative to the control:

(7)	a	Emily wondered who the performer in the concert had <u>imitated</u> for the audience's amusement
	b	Emily wondered which pop star the performer in the concert had <u>imitated</u> for the audience's amusement
	c	Emily wondered whether the performer in the concert had <u>imitated</u> a popstar for the audience's amusement

Furthermore, P600 distribution is present in the posterior areas. According to the authors, the P600 observed at the verb would reflect (i) the increased structural processing that occurs at that position in the *Wh*–fronting conditions, and (ii) the demands on working memory of linking the *Wh*-phrase to the verb across several intervening words. Although both *Which-N* (7b) and *Who* (7a) conditions displayed a posterior, positive component at the embedded verb contrasting clearly with (7c), the author found some small differences between the *Which-N* and *Who* conditions. First, the positivity was larger for the *Which*-N condition than the *Who*-condition. In addition, the scalp distribution for the two conditions was different. This suggests that although both conditions elicit a positivity at the verb, the generators underlying are not completely the same in each case. The difference between *Who* and *Which* + N, however, is outside the scope of the present study.

On this line, Fiebach et al. (2002) investigated a number of different types of indirect *Wh*-questions in German, comparing canonical subject-initial word order and non-canonical object initial word order (as in 8a-b). They also varied the distance between the *Wh*-word and the following argument DP by inserting some additional words between them (the words in square brackets in 8a-b). They found that the P600 might reflect the relative difficulty of integrating a displaced constituent into the verb phrase, which varies depending on the distance from the loci of extraction.

(8)	a. *Who*-subject			
	Thomas fragt sich,	**wer** _	am Dienstag [nachmittag nach dem Unfall]	
	Thomas wonder,	*Who(subject)*	*on Tuesday [afternoon after the accident]*	
	den Doktor verstaendigt hat.			
	the(object) doctor informed has.			
	b. *Who*-object			
	Thomas fragt sich	**wen**	am Dienstag	[nachmittag nach dem Unfall]
	Thomas wonders	*Who(object)*	*on Tuesday*	*[afternoon after the accident]*
	der Doktor _			verstaendigt hat.
		the (subject) doctor		*informed has.*

Intriguingly, when P600s are investigated in shorter *Wh*-dependencies in English vs. longer ones (Phillips et al., 2005), earlier P600 onset latencies are found in the short dependency conditions. The authors proposed that the latency difference reflects the longer time needed to retrieve a more distant *Wh*-phrase from memory. More interestingly, Gouvea et al. (2010) used a within-subjects design with maximally similar materials to enquire syntactic violations, garden path sentences, and the completion of long-distance *Wh*-dependencies in English. As for the latter condition, they used a fronted prepositional *Wh*-phrase (e.g., 'to whom') at the beginning of a relative clause (9).

(9)	The patient met the doctor *to whom* the nurse with the white dress showed the chart during the meeting.

The scalp topography of the P600 was very similar across conditions highlighting a common mechanism underlying this response. However, the P600 elicited by *Wh*-dependencies was smaller than that generated by garden paths, ungrammaticality, and the *Wh*–dependency conditions in previous studies (e.g., Kaan et al., 2000; Phillips et al., 2005). According to the authors, this difference is due to the fact that participants were reading dative-marked *Wh*-phrases (e.g., to whom) that included a fronted preposition. This made it possible to closely match the words following the critical verb, but also meant that participants were already able to identify the case and thematic properties of the *Wh*–phrase at the beginning of the relative clause. This reduced the number of structural relations that needed to be constructed at the verb position, and therefore reduced the amplitude and duration of the P600 relative to earlier studies.

Regarding interrogative sentences, Felser et al. (2003) studied *Which*-object questions in German and showed that sentences containing the *Wh*-dependency at the sub-categorizing verb produced a parietal positivity (P600) rather than a long object topicalization structure. According to the authors, these results support the claim that separable parsing processes are involved in processing syntactic dependencies. Specifically, working memory-based processes are reflected in a LAN for both long object topicalization and *Which-object* dependencies more so than control sentences (short object topicalization sentences). There is also relative difficulty in integrating the filler with its sub-categorizer reflected in a P600 for *Which-object* sentences. According to previous findings on Japanese (Ueno & Kluender, 2009), P600 effects have not been found for direct or indirect *Wh*- questions in Mandarin, as in these languages the Wh-element remains *in situ* (unmoved), hence there is no need to reactivate the *Wh*-filler (Lo & Brennan, 2021).

Within this perspective, an interesting case is represented by *Who*-questions where the Italian interrogative pronoun "chi" has an immediate number disambiguation on the verb following it – singular when "chi" is the subject, plural when it is the object (Penolazzi et al., 2005) – as exemplified in (10).

(10)	a.	Chi bacia (sing.)	i nonni (obj.)	con affetto?
		Who kisses	the grandparents	with affection?
		'Who kisses the grandparents affectionately?'		
	b.	Chi baciano (pl.)	i nonni (subj.)	con affetto?
		Who kiss	the grandparents	with affection?
		'Who do the grandparents kiss affectionately?'		

The results showed a P300 effect (frontally and left distributed) at the verb suggesting quick detection of the syntactic anomaly caused by the ambiguity, corresponding to the first stage of the sentence reanalysis process (see also Osterhout & Holcomb, 1992). In parallel, the verb elicited the classical P600 interpreted in the filler-gap dependencies as a mark of the cost of additional, syntactically guided, revision processes, which are triggered by the perception of an anomaly in the object-first structure.

Let us note that these kinds of revision processes (modulating the N400 and the P600) were also found during the integration of new and given definite Determiner Phrases and the comprehension of the same phrase depending on an inferentially licensed relationship (Burkhardt, 2006, 2007), as well as during presupposition processing in conditions of satisfaction and accommodation (Domaneschi et al., 2018; Jouravlev et al., 2016). Overall, these findings support the view that the P600 reflects discourse-related effects in addition to morpho-syntactic processes (cf. Section 5. for an in-depth discussion).

1.4. The present study

While Penolazzi and colleagues focused their investigation on the temporary syntactic ambiguity of the Italian pronoun *Who-* (subject/object), here we investigate the parsing of Italian *Which-N* questions, where the extraction site could be either the subject or the object noun phrase. For this second type of question, the only data available on Italian come from an eye tracking study performed by Pagliarini et al. (2013), who showed that object *Wh-N* (both *Which-N* and bare *Wh-*) involves higher reading times than subject *Wh-N*. As for the difference between subject and object *Which-N,* no ERP study has, to our knowledge, been previously conducted.

Thus, an issue that has yet to be settled is whether subject and object *Which*-N are differently parsed in interrogative sentences in both behavioral and electrophysiological experiments. In order to address this issue, we based our theoretical assumptions on the Minimal Chain Principle (De Vincenzi, 1991a, 1991b) and verified them using a self-paced reading paradigm together with Event Related Potentials. Accordingly, we assume that unnecessary chain members are avoided and predict that the chain is closed at the earliest point grammatically possible. Therefore, at the behavioral level we hypothesize that the *Which*-N subject will be immediately interpreted and the reading time will be faster than for the *Which*-N object (as in Pagliarini et al. (2013); in fact, only in the case of the *Which*-N object is a reanalysis necessary. We also expect a longer reading time on the post-verbal noun if it has to be taken as the subject. The parser would close the chain immediately after the verb (since Italian allows null subjects and the sentence can be subject-less). In addition, the presence of an overt postverbal subject requires the parser to postulate another multi-membered chain expressed by the dotted line in (5), connecting the postverbal subject DP with the null pronominal *pro*.[3] In parallel, as the P600 seems to reflect reanalysis and reparation processes during structure building (in this case the difficulty to integrate the filler with its sub-categorizer), we hypothesize that, accordingly with the Minimal Chain Principle, the *Which*-N object will show more integration difficulty than the *Which*-N subject and this should be reflected by P600 modulations.

2. Experiment 1: self-paced reading

2.1. Methods

2.1.1. Participants

Twenty-four students of the University of Chieti participated in the experiment (20 females, age-range 19–35 years). All were native Italian speakers, were right-handed (Edinburgh handedness inventory, Oldfield, 1971), and had normal or corrected-to-normal vision. None reported use of psychoactive or vasoactive medication. The participants' written informed consent was obtained according to the Declaration of Helsinki after approval by the ethical committee of the IRCCS Santa Lucia Foundation.

2.1.2. Materials

A total of 60 pairs of sentences were constructed, seven words (W1–W7) in length, with the interrogative pronoun "*Quale*-Noun", as in (4a) and (4b). The critical word, the verb, was always in the second position. The verb was transitive, in the present tense, with 3rd person singular or 3rd person plural inflection. The *Quale-N* and the post-verbal nouns had a human referent. In half of the sentences, the *Quale-noun* was singular, and the post-verbal noun was plural. In the other half, the *Quale-N* was plural (precisely, *Quali-N,* with the inflection marked for the plural noun) and the post-verbal noun was singular. Given that every sentence had a singular and a plural form, both *Quale-N* and the post-verbal noun were the subject an equal number of times.

(11)	a.	Quale ballerina guarda i ragazzi con malizia?
Which dancer looks at the boys with malice?		
	b.	Quale ballerina guardano i ragazzi con malizia?
Which dancer look at the boys with malice? i.e., Which dancer do the boys look at with malice?		

The experimental sentences were counterbalanced across two stimulus lists, so that each participant saw only one version of each

[3] A delay in closing the *Wh*-chain might also be expected because the postverbal subject may work as an intervening element in the object chain, in the terms of Friedmann et al. (2009), as mentioned above (see Guasti, 2017).

pair. Each list also contained 120 sentences with different structures as fillers. All experimental sentences are listed in the Appendix.

2.2. Procedure

We used a self-paced reading paradigm[4] to examine how the difference between subject and object *Wh*-N taps into cognitive processes (Keating & Jegerski, 2015; Marsden et al., 2018). Every sentence was presented word-by-word at the center of the screen and the participant pushed a button to see the following word. Reading Times (RTs) were measured from the word onset to the next press of the button. The task of the participant was to read every word in the sentence as quickly as possible, and to answer a comprehension question about some (one out of ten) sentences. To familiarize the participant with the task, a practice session was run. This consisted of 10 sentences, resembling the experimental and filler items.

2.3. Results

2.3.1. Accuracy

Participants gave the correct response to 87% of the *Which*-subject questions and 84% of the *Which*-object-comprehension questions, indicating that they performed the reading task carefully.

2.4. Reading times

Mean reading times (RTs) and the standard deviation for each word are displayed in Fig. 1. RTs lower than 100 ms and higher than 2500 ms (3.52%) were excluded from further analysis.

A repeated measure analysis of variance (ANOVA) was run on the RTs with the Grammatical Role at two levels (*Which-subject, Which-object*) as independent variable separately for both participants (F1) and items (F2) and from the third to the seventh word position (W3-7) since the RT for W1-2 (the pronoun "Quale" and the noun) was almost identical. Bonferroni correction was applied to p-values (between square brackets) to compensate for multiple comparisons. On the disambiguating W3 (the verb) a significant RT increase for the *Which-object* condition was observed ($F1_{1,23} = 8.39$, $p = 0.008$ [$p = 0.043$], MSE = 14.667 and $F2_{1,59} = 7.03$, $p = 0.010$ [$p = 0.051$], MSE = 14.621) when the verb differing in number was read. The effect of the Grammatical Role was also observed for participants on the W4 (the determiner), as reading in the *Quale-object* condition was slowed down ($F1_{1,23} = 20.26$, $p < 0.001$ [$p > 0.001$], MSE = 4.280). This effect was not significant after Bonferroni correction for the items analysis ($F2_{1,59} = 5.39$, $p = 0.024$ [$p = 0.118$], MSE = 8.113) and for the W5 (second noun) reading ($F1_{1,23} = 4.42$, $p = 0.047$ [$p = 0.232$], MSE = 16.040 and $F2_{1,59} = 6.07$, $p = 0.017$ [$p = 0.066$], MSE = 12.041). For W6–W7 the effect of the Grammatical Role was non-significant ($F < 1$).

3. Experiment 2: ERP study

3.1. Method

3.1.1. Participants

Twenty-one students of the University of Rome Foro Italico participated in the experiment (15 female, mean age 25.2, range 18–35 years). All were native Italian speakers, were right-handed (Edinburgh handedness inventory, Oldfield, 1971), and had normal or corrected-to-normal vision. None had neurological or psychiatric disorders, had experienced neurological trauma, or reported the use of alcohol or other recreational drugs. The participants' written informed consent was obtained according to the Declaration of Helsinki after approval by the ethical committee of the IRCCS Santa Lucia Foundation.

3.1.2. Materials

The stimuli were exactly the same materials as in the reading time study and consisted of a set of 60 sentences constructed including 30 incorrect items and 30 regular sentences. However, unlike the reading time task, the reading was not self-paced, but each word appeared sequentially, not requiring manual intervention. Given that stimulus size is an important variable in ERP studies, each letter height subtended a visual angle of 2° (the indication of the letter size is useful in order to coherently replicate the study).

Stimuli were counterbalanced across two stimulus lists such that each one contained only one version of each experimental pair. Sentences appeared on a PC monitor one word at the time with each word appearing in the center of the screen for 300 ms. The inter-

[4] The self-paced reading paradigm is used to elicit behavioural data, similarly to eye-tracking, **except that it** is constrained by a word-by-word presentation that yields an extremely unnatural reading and can therefore alter the reading process itself. On the other hand, word-by-word reading could be highly informative for experimental purposes because it does not offer options for readers to skip any word (Just et al., 1982) and it reduces the number of compensatory reading strategies available to the readers (Witzel et al., 2012; Jegerski, 2014). Moreover, studies comparing self-paced reading with eye tracking measures report contrasting evidence. While some studies show a lower sensitivity of self-paced reading in tracking the time course of processing mechanisms (Klimek-Jankowska et al., 2018), other studies do not find substantial differences in the capability to detect processing difficulties (Traxler et al., 2002; Koornneef & Van Berkum, 2006; Lyu et al., 2020). Although caution is needed in interpreting self-paced reading data (Miller, 2015; Müller & Mari, 2021), they could nevertheless shed light on the nature of specific processes underlying sentence comprehension.

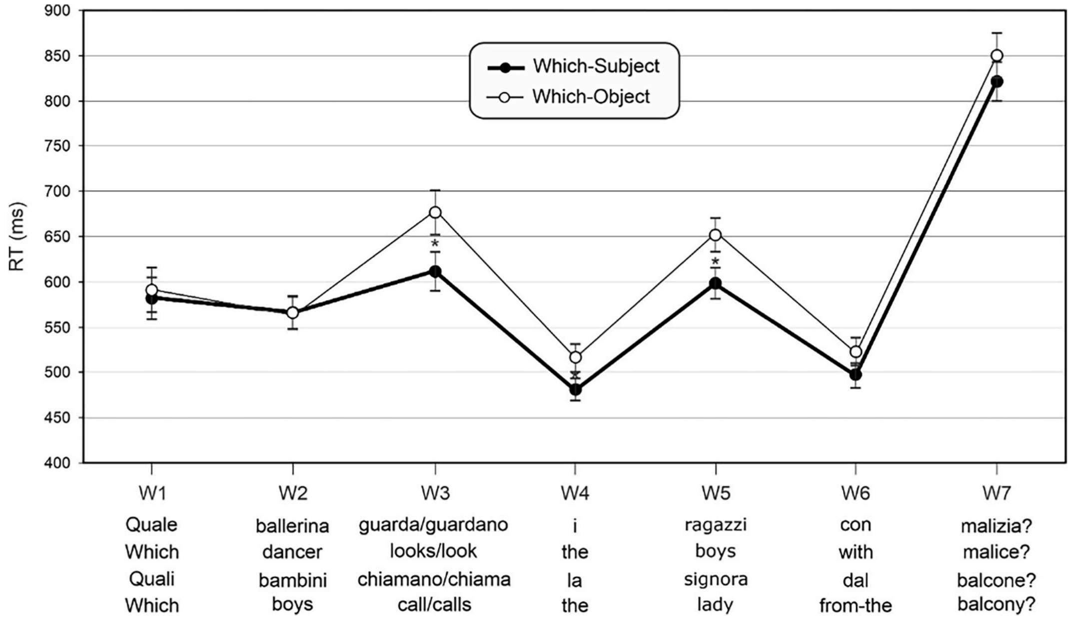


Fig. 1. Average reading times for the *Which*-subject extraction sentences (empty circles) and the *Which*-object extraction sentences (filled circles). Vertical bars denote the standard deviation. *Significant effects.

stimulus interval was 300 ms and the inter-trial interval was 1500 ms. The background luminance (22 cd/m^2) was uniform in intensity and equal to the mean luminance of the stimuli.

3.1.3. Procedure

Subjects were comfortably seated in a dimly lit sound-attenuated and electrically shielded room while stimuli were presented in binocular vision at a viewing distance of 100 cm. Subjects were trained to maintain stable fixation on the center of the screen throughout stimulus presentation. Each run lasted 2 min followed by a 60 s rest period, with longer breaks interspersed. A total of 10 runs were carried out. Subjects were asked to read a series of sentences, some of which were strange or incorrect. They were told to read as rapidly as possible while maintaining good comprehension. To ensure that participants paid attention to the sentences there were some comprehension questions randomly seeded (one every ten sentences) in each experimental run. They had to answer a question (by pressing a "yes" or a "no" button) about the content of the last sentence. A brief practice session was run to familiarize the participants with the task. The subjects were given feedback on their ability to maintain fixation.

3.1.4. Electrophysiological recording

The EEG was recorded from 64 electrodes placed according to the 10-10 system montage (Fp1, Fp2, AF3, AFz, AF4, F7, F3, Fz, F4, F8, FC5, FC3, FC1, FCz, FC2, FC4, FC6, T7, C5, C3, C1, Cz, C2, C4, C6, T8, M2, TP7, CP5, CP3, CP1, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, PO1, POz, PO2, PO4, PO8, O1, Oz, O2, I5, I3, Iz, I4, I6, SI3, SIz and SI4). All scalp channels were initially referenced to the left mastoid (M1) and the re-referenced to the M1-M2 average. Horizontal eye movements were monitored with a bipolar recording from electrodes at the left and right outer canthi. Blinks and vertical eye movements were recorded with an electrode below the left eye, which was referenced to site Fp1. The EEG from each electrode site was digitized at 250 Hz with an amplifier bandpass of 0.01–80 Hz including a 50 Hz notch filter and was stored for off-line averaging. Electrical activity was recorded using two BrainAmp 32-channel amplifiers and the Recorder 1.1 software. The EEG was analyzed using the Analyzer 2.1 software (all by BrainProducts GmbH, Munich, Germany). Computerized artifacts rejection was performed prior to signal averaging in order to discard epochs in which deviations in eye position, blinks, or amplifier blocking occurred. Eye movements were the most frequent cause for rejection. Trials rejected for violating these artifact criteria were 10.7% for the which-object condition and 11.4% for the which-subject condition ($t < 1$). A total of 30 trials for each condition (object and subject) were averaged. ERPs locked to the onset of the critical word and extending until the sentence end (3000 ms) were averaged separately according to the conditions. The amplitudes of the different ERP components were measured as peak values within specified windows with respect to the 200 ms pre-stimulus baseline. To further reduce high-frequency noise, the averaged ERPs were low-pass filtered at 20 Hz.

ANOVAs were used to evaluate the effects of violation on the different ERP components comparing ten 60 ms intervals in the P600 range (500–1100 ms after the critical word onset) to the end of the phrase in pools of electrodes. The Bonferroni correction was applied to the results. The significance level was set at $p < 0.05$. Pools of adjacent electrodes at midline, medial–lateral, and lateral sites were treated separately to allow for quantitative analyses of hemispheric differences. For midline sites, two-way ANOVAs were performed, with repeated measures on two levels of Grammatical Role (Which-Object vs. Which-Subject) and four levels of Electrode Position: Frontal (AFz and Fz), Central (FCz and Cz), Parietal (Pz and POz) and Occipital (Oz and Iz). For medial–lateral electrode sites, three-way ANOVAs were performed on two levels of Grammatical Role, two levels of Hemisphere (left, right), and four levels of Electrode Position: Frontal (AF3/4, F3/4), Centro-frontal (FC3/4 and FC1/2), Central (C3/4, C1/2), Centro-parietal (CP3/4 and CP1/2), Parietal (P3/4, P1/2) and Occipital (PO1/2, O1/2). A three-way ANOVA model was also used for analysis of lateral sites, with repeated measures on two levels of Grammatical Role, two levels of Hemisphere, and three levels of electrode position: Frontal (F7/8, FC5/6), Temporal (T7/8 and C5/6) and Temporo-parietal (TP7/8, P7/8). Post-hoc comparisons were based on the Bonferroni test.

3-D topographical mapping and estimation of the cerebral sources of significant effects were carried out using the Brain Electrical Source Analysis system (BESA 2000 v. 5.1, BESA GmbH, Gräfelfing Germany). The BESA algorithm estimates the location and orientation of equivalent dipolar sources by calculating the scalp distribution that would be obtained for a given dipole model (forward solution) and comparing it to the original ERP distribution. Interactive changes in the location and orientation of the dipole sources led to minimization of the residual variance (RV) between the model and the observed spatiotemporal ERP distribution (for more detail see Di Russo et al., 2005). In these calculations, BESA made a realistic approximation of the head with the radius obtained from the average of the subjects (87 mm). The realistic model of the head is an improvement over the classical spherical approximation (Di Russo et al., 2005). A single dipole was fit over a specific latency range (given below) to correspond to the significant difference between the object and subject condition. The reported dipole fits remained consistent as a function of the starting position. A Polhemus FastTrack spatial digitizer (Polhemus Colchester, Vermont, USA) was used to record the three-dimensional coordinates of each electrode and of three fiducial landmarks (the left and right pre-auricular points and the nasion). A computer algorithm was used to calculate the best-fit sphere that encompassed the array of electrode sites and to determine their spherical coordinates. The mean spherical coordinates for each site averaged across all subjects were used for the topographic mapping and source localization procedures. In addition, individual spherical coordinates were related to the corresponding digitized fiducial landmarks and to the fiducial landmarks identified on the standardized finite element model (FEM) of BESA 2000. The standardized FEM was created from an averaged head using 24 individual MRIs in Talairach space. The averaged head is used for the standard MRI displays and shows a 3D brain. Major sulci can be identified. The standardized FEM provides a realistic approximation of the average head.

4. Results

4.1. *ERP waveforms*

The general shape of the potentials was the same in all conditions, consistently with previous reports (De Vincenzi et al., 2003; Kim & Osterhout, 2005; Osterhout & Nicol, 1999). The ERPs included the typical visual C1, P1, N1 and P2 components peaking around 80, 100, 150 and 180 ms, respectively. The C1 component was medially distributed in parietooccipital areas, the P1 and N1 were bilaterally distributed over occipital areas and the P2 was medially distributed peaking over centro-frontal areas. These potentials were followed by a negative-going component peaking anteriorly at around 400 ms (N400). Fig. 2 shows the ERPs obtained for both the Which-Object (black lines) and Which-Subject (red lines) version of the phrase recorder from the central, parietal, and parieto-occipital electrodes. The time zero represents the onset of the critical word (W3) and vertical lines the onset of successive words (W4-7).

The object version became more negative than the subject starting at 740 ms and peaking at 810 ms on electrode C3. In the medial-lateral ANOVA, this negative difference became significant from 740 to 1040 ms as emerging in the interaction of stimulus Grammatical Role x Hemifield x Electrode ($F_{(3,59)} = 4.5$, $p = 0.006$). Bonferroni correct post-hoc tests showed that the effect was significant for the left Central ($p = 0.014$) and Centro-parietal electrodes ($p = 0.015$). No statistical differences were observed in any analysis in the 500–680 ms range or the 1040–1100 ms range ($F > 1$).

4.2. *Topographical mapping and source analysis*

Fig. 3a shows the 3D scalp topography of the violation effect in relevant time windows where differential activity was maximal.

The activity peaking around 800 ms (Fig. 3a) showed a positive radial topography centered on left centro-parietal regions with steeper gradient on the posterior side suggesting a source in left parieto-temporal areas. Spatiotemporal source analysis in the significant time windows was modeled with a single source localized within structures of the left temporal lobe in the superior temporal gyrus in the proximity of Brodmann's area 22. The dipole orientation pointed toward the parasagittal left centro-parietal scalp as visible on both the realistic (Fig. 3b) and schematic views of the brain (Fig. 3c). The source time-course (Fig. 3d) showed a close resemblance to the surface P600 effect. The Talairach x, y, z coordinates (Talairach & Tournoux, 1988) of the modeled dipole were −49, −18 and 8, respectively. This model explained 96.5% of the variance in the 700 and 1000 ms time windows.

5. Discussion

Thanks to behavioral and neurophysiological approaches, we addressed the question whether the *Which-N* subject extraction in interrogative sentences involves less processing costs as compared to the *Which-N* object extraction, in line with the Minimal Chain Principle (De Vincenzi, 1991a).

The results of the reading time experiment showed that participants take longer to read the verb and the postverbal subject in sentences where the *Quale-N* has an object role. Interestingly, wrap-up processes on the final words (after the postverbal subject) of the *Quale-object* do not require extra time. This finding suggests that the parser proceeds in an incremental manner, following grammatical principles and interpreting the interrogative phrase as a subject. Higher reading times for verbs which do not match the preceding Determiner Phrase (DP) are also found in the eye tracking study of Pagliarini et al. (2013) and in other Self-paced Reading paradigm

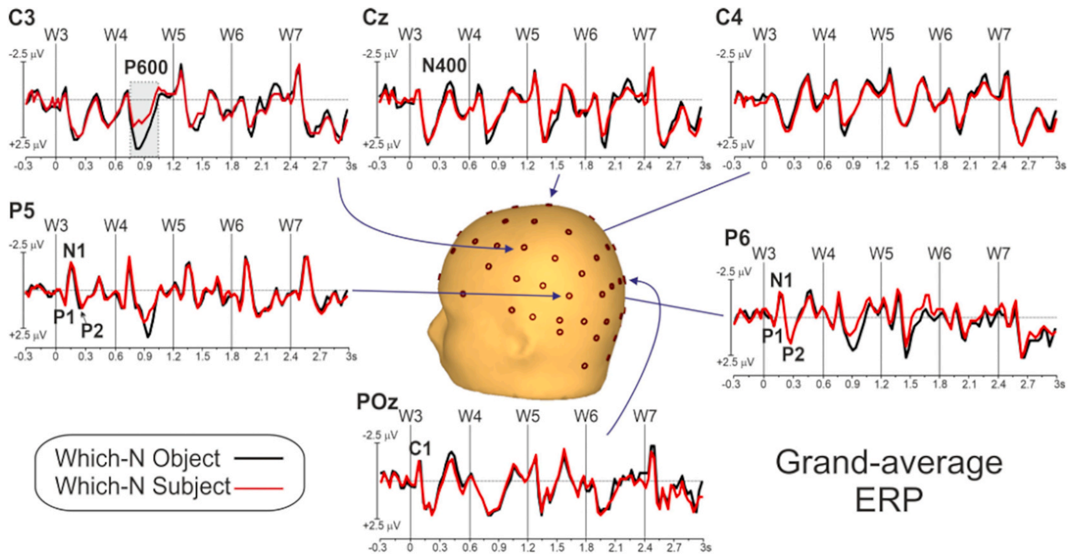


Fig. 2. Grand-average ERP waveforms triggered by the critical word (W3) for *Which*-object extraction (black lines) and *Which*-subject conditions (red lines). Electrodes are shown from selected locations (central, parietal and parieto-occipital) indicated on the 3-D head. Major ERP components are labeled. **The** highlighted time window represents significant effects. **The** vertical dotted line represents word onset.

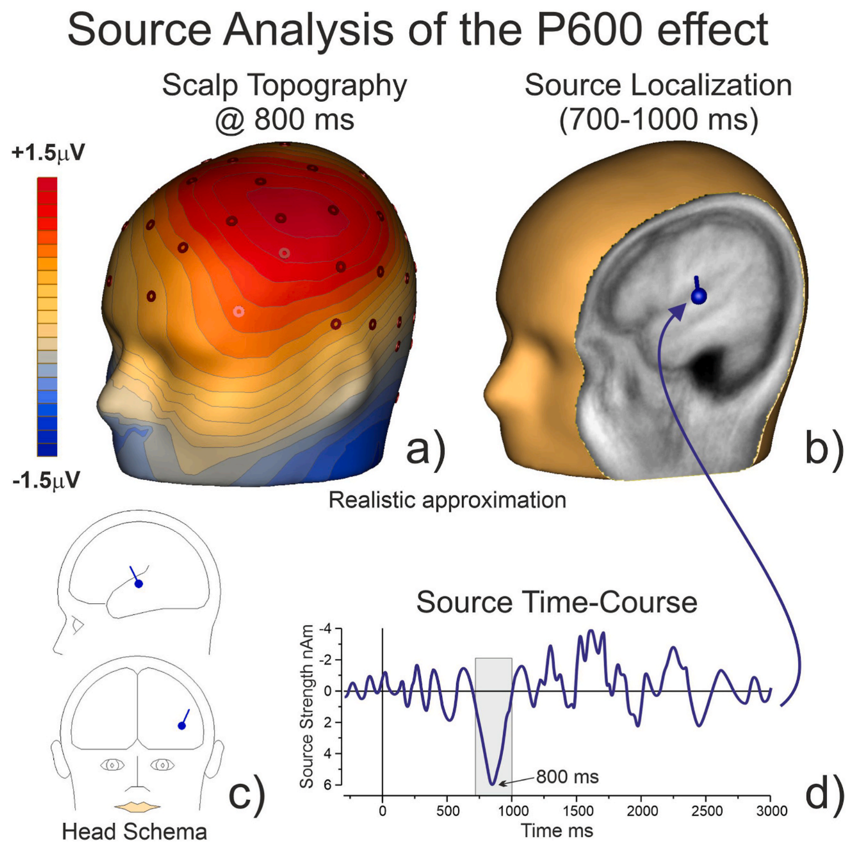


Fig. 3. a) Spline-interpolated 3-D topographic map of the differences between the ERP in response to the 'violate' and the correct condition at 800 ms after the critical word onset. Red areas indicate positive differences between conditions, and blue areas indicate negative differences. Lines represent iso-voltage areas. Electrode positions are marked by circles. b) Shows the source localization and orientation (blue sphere with pointer) of effect from 700 to 1000 ms on a realistic brain. c) Head schema of source position and orientation from left-lateral and frontal view. d) Time-course of cerebral generator strength (dipole moment) peaking at 800 ms.

experiments involving number to disambiguate between subject and object relative clauses (Arosio et al., 2009; Villata & Lorusso, 2020). Crucially, as anticipated above, we also saw a longer reading time on the post-verbal noun when it had to be taken as the subject. We can hypothesize that this is a spillover effect after the reanalysis at the verb region which may take longer (as suggested by Pagliarini et al., 2013). Another option is to account for this finding in the terms of the Minimal Chain Principle: in the *Which-subject* condition, as in (3), (12), the only chain is closed when the verb is parsed. In the *Which-object* condition the parser has to perform a reanalysis at the verb region and a second chain has to be postulated when the postverbal subject is parsed, as in (5), (13). The parser avoids creating an unnecessary chain until new evidence requires it.

In the *Which-subject* condition the *Quale-N* matches the number of the verb, so the chain is closed as soon as the verb is parsed (12).

(12) a. subject which question:

a. subject which question:
Wh DP1$_i$ Q . . . ~~<**Wh DP1**> I$_i$~~ . . . D DP2

In the *Which-object* condition, since the *Quale-N* and the verb mismatch in number, the *Which-subject* interpretation must be dropped and the sentence reanalyzed. However, the parser can still interpret the *Which-N* as the object of the verb in a subjectless sentence, but when the postverbal subject is processed a new chain has to be postulated between the preverbal *pro* and the postverbal subject; this entails a delay in processing.

(13) b. object which question:

b. object which question:
WhDP2Q$_k$.... *pro*$_j$ I DDP1$_j$...... <**Wh DP2**>$_k$

Our findings on the higher reading times with *Which*-N object sentences fully confirm the Dutch data collected by Frazier and Flores D'Arcais (1989), which show faster recognition for *Which*-N subject sentences. Participants had to give a judgment if sentences were acceptable or not. Therefore, the dependent measure was the percent and timing of the correct responses. Sentences with subject extraction were better understood than *Which*-object extraction. Let's notice that the judgement on Dutch sentences is given at the end of the sentence, so we do not know exactly what is happening on-line, that is when the sentence is read word-by-word. However, our result seems to confirm a fast decision, with some penalty on the post-verbal noun.

This work is apparently in contradiction with what has been found by Stowe (1986), who did not find a subject expectation in English. The data that she used include only *Who*- dependencies. The relevant contrasts that she uses are the following (10):

(10)	a.	My brother wanted to know if Ruth will bring us home to Mom at Christmas.
	b.	My brother wanted to know who will bring us home to Mom at Christmas.
	c.	My brother wanted to know who Ruth will bring home to Mom at Christmas.

The dependent measure was reading time word-by-word at the center of the screen. The reading time on the word "Ruth" was not significantly different between (10a) and (10c), providing no evidence that people expect a gap at the subject position (10c). One difference with our data is that we had a stronger disambiguation given by the number mismatch on the verb (cf. De Vincenzi et al., 2003). Another crucial difference is that English does not allow postverbal subjects. In Italian *Wh*-questions, a postverbal subject is not only allowed but obligatory, when the subject is not dropped, as a consequence of the *Wh*-criterion (Rizzi, 1997). So, while the preverbal subject is expected in English (and no differences are found between 10a and 10c), in Italian the preverbal subject is never allowed nor expected in *Which-N* objects: the preverbal position is only available as a trace in *Which-N* subject sentences.

For what concerns the ERP findings, our data confirmed and extended those of previous studies in showing that the P600 reflects the relative difficulty of integrating the fillers with their sub-categorizers, both for direct and indirect interrogative sentences (cf. Fiebach et al., 2002; Phillips et al., 2005; Goueva et al., 2010; Felser et al., 2003; Penolazzi et al., 2005).[5] In fact, we clearly found that the *Wh-* condition (*Quale/Quali)* showed an effect on the positive component of the P600 in the 760–1040 ms interval on the verb, when the *Which*-N has to be taken as object. This effect was limited to the left hemisphere and was localized within the superior temporal gyrus. More in depth, this result also agrees with what Mecklinger et al. (1995) and Friederici et al. (2001) found for ambiguous relative clauses. These authors found a P600 on the last word, the auxiliary verb, that disambiguates the subject or object role of the noun phrase through the number marking on the noun phrases. The P600 is found in the object relative clause sentence, and specifically in reading the auxiliary in the final position (cf. (14)). Although the relative sentences in (14) do not involve *Wh-N* movement, it is relevant for our findings since, as in our case, a subject-object asymmetry disambiguated by the number of the verb involves a P600 for the object condition although both sentences (14a-b) are grammatical.

(14)	a.	Chi-subject Das ist die Direktorin, die die Sekretaerinnen gesucht hat. This is the director, that the secretaries looked for has.
	b.	Chi-object Das ist die Direktorin, die die Sekretaerinnen gesucht haben This is the director, that the secretaries looked for have.

Tanenhaus et al. (1989), showed an immediate effect of lexical verb preference on *Which* assignment, while Frazier and Clifton (1989) found no effect of a verb-preferred assignment on *Who* sentences. As for German, Farke and Felix (1994) found an object extraction preference with which, while Schlesewsky et al. (1999) found a *Who* subject extraction preference. In both languages, the difference is explainable under the assumption that *Who*, but not *Which*, enters into a long-distance relationship, subject to syntactic requirements. Therefore, an immediate subject extraction is found only for *Who*, with no intervention of verb preferences.

Finally, a multi-word sustained anterior negativity has been observed in object relative clauses (which involve a long-distance dependency) with respect to the subject relative clause control condition (King & Kutas, 1995; see also Fiebach et al., 2002). This effect seems to begin shortly after the head of the relative clause and continues to the verb of the relative clause four words later. It has been argued that the sustained negativity reflects the working memory load of holding an incomplete syntactic dependency in mind, reflecting length-sensitive processes. However, Phillips et al. (2005), for relative clauses, observed that the words following the relative pronouns do not contribute to the greater negativity effect. In fact, the last four words of the dependency make no contribution at all to the sustained negativity, contrary to the predictions that sustained negativity reflects gradually increasing memory costs associated with incomplete *Wh*-dependency. Accordingly, it seems that the sustained anterior negativity does not provide evidence that the cost of storing the *Wh*-phrase in memory increases over the course of the dependency, or simply that the sustained anterior negativity is an insufficiently sensitive measure of working memory load in sentence processing. Our data confirm these hypotheses, as in subject and object *Wh-N* interrogative sentence parsing we did not find any end-of-sentence cost, such as a long negativity. These results are perfectly in line with our syntactic parsing model (Frazier & Rayner, 1982; De Vincenzi, 1991a, 1991b): it, in fact, predict no extra cost of reading the sentence, which is resolved very quickly (see similar results in Frazier et al., 2015 for English *Wh*-filler-gap dependencies and reflexive-antecedent dependencies, and in Lo & Brennan, 2021 for Mandarin Wh-questions).

This issue leads us to discuss a general problem linked to the intrinsic nature of the P600. Actually, we know that ungrammaticality, garden-paths, and the completion of long-distance dependencies generate the P600 response component. However, documentation increasingly shows that the P600 may vary along a number of different parameters, including latency, duration, amplitude, and scalp topography. This multi-dimensional property of the P600 suggests that, on the one hand, this component reflects a cohesive set of processes, on the other hand different parsing computations may modulate these processes. According to Gouvea et al. (2010), the P600 may be understood as reflecting a common set of processes that occur on a just-in-time basis, beginning as soon as sufficient information has been gathered to initiate the processes. Under this view, when the P600 occurs at different latencies it reflects the same underlying processes, with latency variation displaying the time needed to complete the processes that trigger it (see also Friederici et al., 2001; Hagoort, 2003). According to this line of argument, it follows that while the latency of the P600 should be an index of the time needed for retrieval of the elements that participate in structural relations, the duration and amplitude of the P600 should be a function of the structure building processes themselves. Conversely, the latency of the P600 should be an index of the time needed for

[5] As an anonymous reviewer pointed out, the P600 found with *Which*-objects might be related to the lower frequency of *Which*-object sentences. However, to our knowledge no data of a higher frequency of subject over object interrogative clauses has been uncontroversially attested in spontaneous Italian speech. An elicited production study of Wh-questions in Italian showed that adults produce almost the same number of correct subject which-questions as object which-questions (around 83%) (Guasti, Branchini, & Arosio, 2012). This may indicate that subject which-questions are not more frequently produced than object ones. Studies on the spontaneous speech of English, like the one of Stromswold (1995), showed that children start to produce object interrogatives before subject interrogatives while in adult speech subject questions are more common. The possibly higher frequency of subject questions, however, would not logically imply a lack of production or comprehension (and a P600) of object questions which are in general very common and productive as studies like Stromswold's (1995) suggest. Furthermore, in our data the P600 effect is time related to the processing of verbal morphology in matching/mismatching the preverbal Which-N.

retrieval of the elements that participate in structural relations. As a consequence, different structural and lexical manipulations should impact the P600 differently. The manipulations that impact retrieval processes should change the latency of the P600, whereas those that impact the number and type of syntactic relations should change the its amplitude and/or duration (Phillips et al., 2005; Gouvea et al., 2010).

Accordingly, we found that the parsing related to the *Wh*-object interrogative sentences elicited wider P600 amplitude than the *Wh*-subject interrogative sentences. Thus, we added one more piece of evidence that the amplitude of the P600 is concerned with the required structure building operations. In fact, the P600 we found may be considered a response to a sort of very brief garden-path and its structural reanalysis, as already highlighted by previous studies (cf. for instance Gouvea et al., 2010). In particular, it might reflect the cost of *revising* a previously selected interpretation. This hypothesis is also compatible with our source data (cf. Fig. 3) showing that the P600 is generated in the left superior temporal gyrus (Brodmann's area 22). In fact, this cortical area seems functionally dedicated to controlling semantic operations correlated with processes of sentential integration (cf. Friederici et al., 2003).

As for the frequency of subject vs object interrogative sentences, one might expect that the P600 in our data could be related to the low frequency of object interrogative constructions in spontaneous speech. To our knowledge, no studies have been performed on the relation between the relative frequency of subject-object interrogatives and their processing. However, studies on the acquisition of English interrogative sentences (Stromswold, 1995) show that there is no correlation between frequency in adult speech and either the order of acquisition of subject and object questions or the percentage of errors found in children's spontaneous production of interrogatives. In the present study the P600 seems to be related to the processing of verbal morphology matching/mismatching the number of the Which-N element. Nonetheless, we are aware that further studies are needed to address the issues in the P600 which can be modulated by frequency effects.

6. Conclusions and further remarks

In summary, we reported data congruent with a model of sentence processing that directly employs syntactic analysis according to the Minimal Chain Principle. Behavioral and ERP data have shown that *Which-N* dependencies are quickly resolved at the verb phrase and at the subsequent noun phrase.

Considering that the noun phrase has to be taken either as an object or a subject, only in the first case does the presence of multiple chains involve a delay in the processing of the sentence as predicted by the Minimal Chain Principle. Actually, the ERP responses give rise to a P600 on the *Wh*-object structure (760–1040). This component is limited to the left hemisphere and localized in the superior temporal gyrus, where semantic and syntactic computations are integrated. This evidence suggests that the P600 reflects the cost necessary to reanalyze structure building operations as in the case of garden-path sentences. Crucially, this process does not require any extra cost of reading the sentence, as suggested by the absence of any long negativity.

Altogether, these findings agree with and strengthen previous data suggesting that the P600 amplitude involves multi-dimensional processes controlling operations such as prediction, retrieval, revising, and structure building needed to create syntactic structures. In particular, we provided further evidence on the hypothesis that the amplitude and duration of the P600 is a function of the assembly (and disassembly) of syntactic relations.

Enlarging the focus of our analysis, we should note that offline comprehension and online (ERP) sentence processing (in particular the P600) may be modulated by both the expectations of the speaker and the type of context information (Burmester et al., 2014); note that Wh-elements are focal (new information) elements since they compete for the same position with focus (as observed in Rizzi, 1997) in direct questions. Indeed, predictions seem to be intrinsically related to semantic and syntactic levels (Martorell, 2018). Thus, we cannot exclude that our behavioral and ERP findings have also been influenced by the fact that object phrases are less expected to appear in the periphery of a sentence. In fact, Burmester et al. (2014) found that the presence of an appropriate context facilitates the comprehension of object-initial German sentences reducing late positivity ERP effects.

Finally, neurophysiological investigation of sentence processing is becoming increasingly focused on oscillatory-based analysis. This perspective might provide a new window on the dynamics of the coupling and uncoupling of functional networks involved in cognitive processing (Varela et al., 2001). How these two measures of the neural correlates underlying cognitive processes relate to one another remains unclear. Given the current evidence, the most systematic effects of morpho-syntactic processing seem to emerge as power decreases in the alpha and/or beta bands involving syntactic category, verb tense, and subject-verb agreement (cf. Grimaldi, 2018; Martorell et al., i-an press). However, investigations on interrogative sentences are lacking and definitive conclusions on how to interpret these (as other) oscillatory rhythms are still premature. In the future, we need to combine different approaches, theoretical models, and analyses in order to reach a better understanding of the neural dynamics that underpin sentence processing in the brain.

Credit author statement

Mirko Grimaldi: Conceptualization, Methodology, Supervision, Validation, Writing - original draft, Writing – review & editing.
Marica De Vincenzi: Conceptualization, Methodology, Validation, Investigation, Writing - original draft.
Paolo Lorusso: Conceptualization, Validation, Writing - original draft, Writing – review & editing.
Francesco Di Russo: Methodology, Software, Formal analysis, Visualization, Writing - original draft, Writing – review & editing.
Rosalia Di Matteo: Methodology, Investigation, Writing – original draft e Writing – review & editing.
Luigi Rizzi: Conceptualization, Methodology, Supervision, Validation, Writing - original draft, Writing – review & editing.
Maria Teresa Guasti: Conceptualization, Methodology, Supervision, Validation, Writing - original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that the research was conducted in the absence of any competing financial, commercial and nonfinancial interests that might be perceived to influence the results and/or discussion reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2023.101154.

References

Arosio, F., Adani, F., & Guasti, M. T. (2009). Grammatical features in the comprehension of Italian relative clases by children. In J. M. Brucart, A. Gavarró, & J. Solà (Eds.), *Merging features: Computation, interpretation, and acquisition* (pp. 138–155). Oxford: Oxford University Press.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language, 98*(2), 159–168. https://doi.org/10.1016/j.bandl.2006.04.005

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport, 18*(17), 1851–1854. https://doi.org/10.1097/WNR.0b013e3282f1a999

Burmester, J., Spalek, K., & Wartenburger, I. (2014). Context updating during sentence comprehension: The effect of aboutness topic. *Brain and Language, 137*, 62–76. https://doi.org/10.1016/j.bandl.2014.08.001

Burzio, L. (1986). *Italian syntax: A government-binding approach*. Dordrecht, Holland: D. Reidel publishing Company.

Chomsky, N. (1981). *Lectures on government and binding. The Pisa lectures*. Dordrecht and Cinnaminson: Foris Publications.

De Vincenzi, M. (1991a). *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer Academic Publishers.

De Vincenzi, M. (1991b). Processing of *wh*-dependencies in a null subject language: Referential and non-referential whs. *University of Massachusetts Occasional Papers in Linguistics, 17*(2), Article 4.

De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., & Ciccarelli, L. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language, 85*(2), 280–296. https://doi.org/10.1016/S0093-934X(03)00055-5

Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D., & Hillyard, S. A. (2005). Identification of the neural sources of the pattern-reversal VEP. *NeuroImage, 24*, 874–886. https://doi.org/10.1016/j.neuroimage.2004.09.029

Domaneschi, F., Canal, P., Masia, V., Lombardi Vallauri, E., & Bambini, V. (2018). N400 and P600 modulation in presupposition accommodation: The effect of different trigger types. *Journal of Neurolinguistics, 45*, 13–35. https://doi.org/10.1016/j.jneuroling.2017.08.002

Farke, H., & Felix, S. W. (1994). Subjekt-objektasymmetrien in der Sprachverarbeitung. In S. W. Felix, C. Habel, & G. Rickheit (Eds.), *Kognitive Linguistik. Repräsentationen und Prozesse. Opladen.* Westdeutscher Verlag.

Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and Wh -movement in German. *Brain and Language, 87*(3), 345–354. https://doi.org/10.1016/S0093-934X(03)00135-4

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language, 47*(2), 250–272. https://doi.org/10.1016/S0749-596X(02)00004-9

Frazier, M., Ackerman, L., Baumann, P., Potter, D., & Yoshida, M. (2015). Wh-filler-gap dependency formation guides reflexive antecedent search. *Frontiers in Psychology, 6*, Article e1504. https://doi.org/10.3389/fpsyg.2015.01504

Frazier, L., & Clifton, C., Jr. (1989). Successive cyclicity in the grammar and the parser. *Language & Cognitive Processes, 4*, 93–126.

Frazier, L., & Flores D'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language, 28*(3), 331–344. https://doi.org/10.1016/0749-596X(89)90037-5

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84. https://doi.org/10.1016/s1364-6613(00)01839-8

Friederici, A. D. (2017). Neurobiology of syntax as the core of human language. *Biolinguistics, 11*(Special Issue), 325–337. https://doi.org/10.5964/bioling.9093

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage, 20*(1), 8–17. https://doi.org/10.1016/j.neuroimage.2003.09.003

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*, 305–323. https://doi.org/10.1016/S0926-6410(00)00065-3

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*(3), 183–192. https://doi.org/10.1016/0926-6410(93)90026-2

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research, 1146*, 50–58. https://doi.org/10.1016/j.brainres.2006.08.038

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua, 119*, 67–88. https://doi.org/10.1016/j.lingua.2008.09.002

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of wh questions in three subtypes of SLI. *Lingua, 121*, 367–382. https://doi.org/10.1016/j.lingua.2010.10.004

Gouvêa, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language & Cognitive Processes, 25*(2), 149–188. https://doi.org/10.1080/01690960902965951

Grillo, N. (2005). Minimality effects in agrammatic comprehension. In S. Blaho, L. Vicente, & E. Schoorlemmer (Eds.), *Proceedings of console XIII* (pp. 106–120). University of Leiden. ISSN: 1574-499X.

Grillo, N. (2008). *Generalized minimality: Syntactic underspecification in broca's aphasia*. Doctoral Dissertation, University of Utrecht, The Netherlands: Landelijke Onderzoekschool Taalwetenschap (LOT).

Guasti, M. T. (2017). *Language acquisition. The growth of grammar*. Cambridge: MIT Press.

Guasti, M. T., Branchini, C., & Arosio, F. (2012). Interference in the production of Italian subject and object wh-questions. *Applied PsychoLinguistics, 33*, 185–223. Published online June 2011.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*(6), 883–899. https://doi.org/10.1162/089892903322370807

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30*(1), 47–61. https://doi.org/10.1111/j.1469-8986.1993.tb03204.x

Jegerski, J. (2014). Self-paced reading. In J. Jegerski, & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge.

Jouravlev, O., Stearns, L., Bergen, L., Eddy, M., Gibson, E., & Fedorenko, E. (2016). Processing temporal presuppositions: An event-related potential study. *Language, Cognition and Neuroscience, 31*(10), 1245–1256. https://doi.org/10.1080/23273798.2016.1209531

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language & Cognitive Processes, 15*(2), 159–201. https://doi.org/10.1080/016909600386084

Keating, G., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition, 37*(1), 1–32. https://doi.org/10.1017/S0272263114000187

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205–225. https://doi.org/10.1016/j.jml.2004.10.002

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. https://doi.org/10.1162/jocn.1995.7.3.376

Klimek-Jankowska, D., Czypionka, A., Witkowski, W., & Błaszczak, J. (2018). The time course of processing perfective and imperfective aspect in Polish. *Acta Linguistica Academica Acta Linguistica Hungarica, 65*(2–3), 293–351. https://doi.org/10.1556/2062.2018.65.2-3.4

Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language, 54*(4), 445–465. https://doi.org/10.1016/j.jml.2005.12.003

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego: Academic Press.

Lo, C.-W., & Brennan, J. (2021). EEG correlates of long-distance dependency formation in Mandarin wh-questions. *Frontiers in Human Neuroscience, 15*, Article e591613. https://doi:10.3389/fnhum.2021.591613.

Lyu, S., Jung-yueh, T., & Chien-Jer, C. L. (2020). Processing plausibility in concessive and causal relations: Evidence from self-paced reading and eye-tracking. *Discourse Processes, 57*(4), 320–342. https://doi.org/10.1080/0163853X.2019.1680089

Marsden, E., Thompson, S., & Polinsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied PsychoLinguistics, 39*(5), 861–904. https://doi:10.1017/S0142716418000036.

Martorell, J. (2018). Merging generative linguistics and psycholinguistics. *Frontiers in Psychology, 9*, 2283. https://doi.org/10.3389/fpsyg.2018.02283

Martorell, J., Morucci, P., Mancini, S., Molinaro, N. (2023). Sentence processing: How words generate syntactic structures in the brain. In M. Grimaldi, Y. Shtyrov, B. Elvira (Eds.), Language electrified: Principles, methods and future perspectives of investigation, Amsterdam-New York, Springer Nature.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition, 23*, 477–494. https://doi.org/10.3758/BF03197249

Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychology, 50*(1), 31–42. https://doi.org/10.1080/00461520.2015.1004068

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47*(8), 908–930. https://doi.org/10.1016/j.cortex.2011.02.019

Müller, M. L., & Mari, M. A. (2021). Definite descriptions in the light of the comprehension vs. Acceptance distinction: Comparing self-paced reading with eye-tracking measures. *Frontiers in Communication, 6*, 634362. https://doi.org/10.3389/fcomm.2021.634362

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*(2), 151–165. https://doi.org/10.1162/jocn.1991.3.2.151

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension. In M. D. Rugg, & M. G. H. Coles (Eds.), *Oxford psychology series, No. 25. Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171–215). Oxford: Oxford University Press.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language & Cognitive Processes, 14*(3), 283–317. https://doi.org/10.1080/016909699386310

Pagliarini, E., Vernice, M., & Guasti, M. T. (2013). The processing Wh-questions in Italian: Evidence from an eye-tracking study. In *Proceedings of AMLAP 2013: Architecture and mechanism of language processing*. University of Marseille, 2nd- 4th September 2013 https://www.dce.unimore.it/site/home/dipartimento/archivio-eventi/igg39/program/documento900-025987.html.

Penolazzi, B., De Vincenzi, M., Angrilli, A., & Job, R. (2005). Processing of temporary syntactic ambiguity in Italian "who"-questions: A study with event-related potential. *Neuroscience Letters, 227*, 91–96. https://doi.org/10.1016/j.neulet.2004.11.074

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research, 22*(3), 407–428. https://doi:10.1016/j.cogbrainres.2004.09.012.

Pollock, J. Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry, 20*, 365–424.

Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations. *PLoS One, 9*(5), Article e96840. https://doi.org/10.1371/journal.pone.0096840

Rizzi, L. (1990). *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar: A handbook of generative syntax* (pp. 281–337). Dordrecht: Kluwer.

Rizzi, L. (2004). Locality and left periphery. In A. Belletti (Ed.), *Structures and beyond: The cartographic syntactic structure* (pp. 223–251). Oxford: Oxford University Press.

Rösler, F., Pütz, P., Friederici, A. D., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5*(3), 345–362. https://doi.org/10.1162/jocn.1993.5.3.345

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (1999). Preferences for grammatical functions in the processing of locally ambiguous Wh-questions in German. In B. Hemforth, & L. Konieczny (Eds.), *Cognitive parsing in German* (pp. 65–94). Dordrecht: Kluwer.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language & Cognitive Processes, 1*(3), 227–245. https://doi.org/10.1080/01690968608407062

Stromswold, K. (1995). The acquisition of subject and object WH-questions. *Language Acquisition, 4*, 5–48.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain* (New York, Thieme).

Tanenhaus, M. K., Stowe, L. A., & Carlson, G. (1989). Lexical expectation and pragmatics in parsing filler-gap constructions. In W. D. Marlsen-Wilson (Ed.), *Lexical representation and process* (pp. 311–331). Cambridge MA: M.I.T. Press. Cambridge.

Tanner, D. (2019). Robust neurocognitive individual differences in grammatical agreement processing: A latent variable approach. *Cortex, 111*, 210–237. https://doi:10.1016/j.cortex.2018.10.011.

Tanner, D., Grey, S., & van Hell, J. G. (2017). Dissociating retrieval interference and reanalysis in the P600 during sentence comprehension. *Psychophysiology, 54*(2), 248–259. https://doi.org/10.1111/psyp.12788

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*(4), 530–547. https://doi.org/10.1016/S0749-596X(02)00021-9

Ueno, M., & Kluender, R. (2009). On the processing of Japanese Wh-questions: An ERP study. *Brain Research, 1290*, 63–90. https://doi.org/10.1016/j.brainres.2009.05.084

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

Villata, S., & Lorusso, P. (2020). When initial thematic roles attribution lingers: Evidence for digging-in effects in Italian relative clauses. In V. Torrens (Ed.), *Typical and impaired processing in morphosyntax* (pp. 57–71). Amsterdam: John Benjamins.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research, 41*(2), 105–128. https://doi.org/10.1007/s10936-011-9179-x