UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
DOUTORADO EM LINGUÍSTICA

Bianca Chaves Leite Lignani

EXHAUSTIVITY, LINGUISTIC PREDICTION AND INTEGRATION: IT-CLEFTS AND 'ONLY' SENTENCES

Juiz de Fora
2026

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
DOUTORADO EM LINGUÍSTICA

Bianca Chaves Leite Lignani

EXHAUSTIVITY, LINGUISTIC PREDICTION AND INTEGRATION: IT-CLEFTS AND 'ONLY' SENTENCES

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Doutora em Linguística. Área de concentração: Linguística.

Juiz de Fora
2026

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Lignani, Bianca Chaves Leite.
Exhaustivity, Linguistic Prediction and Integration: : it-clefts and 'only'-sentences / Bianca Chaves Leite Lignani. -- 2026.
193 p. : il.

Orientadora: Aline Alves Fonseca
Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2026.

1. Advérbios focalizadores. 2. Clivagem. 3. Exaustividade. 4. Predição linguística. I. Fonseca, Aline Alves, orient. II. Título.

Bianca Chaves Leite Lignani

EXHAUSTIVITY, LINGUISTIC PREDICTION AND INTEGRATION: IT-CLEFTS AND 'ONLY' SENTENCES

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de doutor em linguística. Área de concentração: linguística.

Aprovada em 27 de abril de 2026.

BANCA EXAMINADORA

Prof(a) Dr(a). Aline Alves Fonseca - Orientadora
Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Maria Cristina Lobo Name
Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Kevin Paterson
University of Leicester

Prof(a) Dr(a). Ascensión Pagán
University of Leicester

Prof(a) Dr(a). Marcus Maia
Universidade Federal do Rio de Janeiro

Juiz de Fora, 17/04/2026.

Documento assinado eletronicamente por **Aline Alves Fonseca**, **Professor(a)**, em 27/04/2026, às 12:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Maria Cristina Lobo Name**, **Professor(a)**, em 06/05/2026, às 13:21, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **MARCUS ANTONIO REZENDE MAIA**, **Usuário Externo**, em 07/05/2026, às 11:54, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ascension Pilar Pagan Camacho**, **Usuário Externo**, em 12/05/2026, às 10:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Kevin Paterson**, **Usuário Externo**, em 13/05/2026, às 11:22, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

AGRADECIMENTOS

Com muita felicidade, cheguei ao final desta trajetória. Agradeço a todas as pessoas que se fizeram presentes durante esta importante etapa da minha vida, compartilhando os momentos felizes e amenizando aqueles que não foram tão felizes assim. Obrigada:

À minha mãe e ao meu pai, por tudo e para sempre. Sou muito afortunada por tê-los como meus maiores ídolos e meus maiores incentivadores, durante toda a minha vida.

Ao meu amado irmão, por demonstrar sempre carinho e cuidado comigo.

Ao meu namorado, por ter pacientemente me ouvido discorrer sobre o tema desta pesquisa em pelo menos cinco situações diferentes - e ainda parecer entusiasmado.

À minha querida orientadora, Professora Aline Fonseca, pelos valiosos conhecimentos compartilhados comigo ao longo das aulas do Programa de Doutorado e ao longo do Estágio em Docência. Principalmente, agradeço a ela por ter me acompanhado nesta nova etapa acadêmica e por tê-lo feito, novamente, com tanta dedicação, carinho e entusiasmo.

Ao Professor Marcos Maia, por ter contribuído tanto – e de forma tão gentil – com a minha tese durante o meu exame de Qualificação e por ter aceitado, novamente, o convite para compor esta banca examinadora. Foi um enorme prazer finalmente tê-lo conhecido e o ver presente na Banca Examinadora.

À Professora Cristina Name, pelas contribuições durante o exame de Qualificação, pelas maravilhosas aulas no Programa de Doutorado e por ter sido sempre tão solícita, acessível e cordial. Agradeço imensamente o seu aceite para compor esta Banca Examinadora.

Às(os) demais professoras(es) do Programa de Pós-Graduação em Linguística, que expandiram meus conhecimentos.

Às professoras e aos estimados colegas e ex-colegas do NEALP, pelos momentos de expansão de ideias e pelas preciosas vivências acadêmicas compartilhadas.

Aos gentis participantes dos meus experimentos, que possibilitaram o avanço da minha pesquisa.

Aos meus amigos e às minhas amigas, agradeço por terem me motivado tanto, sendo compreensivos e pacientes durante as minhas ausências e/ou restrições.

Agradeço também à CAPES e ao CNPq pelo apoio financeiro, sem o qual essa pesquisa não teria sido possível.

O presente trabalho foi realizado, parcialmente, com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) – processo nº 88881.980957/2024-01.

O presente trabalho foi realizado, parcialmente, com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – processo nº 140293/2026-4.

ACKNOWLEDGMENTS

The completion of this study would not have been possible without the expertise and the guidance of Prof. Kevin Paterson from the University of Leicester (UK). I would also like to credit Dr. Ascen Pagán and Dr. Vicky McGowan, from the University of Leicester.

My most sincere thanks to all of you, who were not only excellent mentors but also great hosts during my stay in the UK. This period will always hold very special memories to me.

Finally, “Cheers!” to my dear flatmates and to the great PhD colleagues that I met in Leicester. I hope we can meet again.

ABSTRACT

This research investigates sentences with the adverb ‘only’/‘*só*’ and it-clefts/*clivadas totais*, in English and in Brazilian Portuguese (BP). There is consistent evidence that these focus markers (Kiss, 1998; Paterson; Liversedge; Underwood, 1999; Liversedge *et al*, 2002; Zimmerman, 2004; Filik *et al*, 2009; Carlson, 2015) convey exhaustivity through different processes (Horn, 1981; Rooth, 1992; Krifka, 1999; Drenhaus; Zimmerman; Vasishth, 2011; Carlson, 2014; Lignani; Fonseca, 2023). Exhaustivity is the semantic effect whereby a focused element in a sentence excludes all other alternatives to make that sentence true (Rooth, 1992; Krifka, 1999). Two types of experiments were designed to compare exhaustivity violations for clefts and ‘only’-structures: a) three Likert tests (Likert, 1932) – AJTs (in English, tasks 1 and 2; in BP, task 3); and b) two eye-tracking tasks (one English and one in BP). Our hypotheses were that: 1) exhaustivity violations in ‘only’-sentences/'*só*'-sentences' are less acceptable than in it-clefts/*clivadas totais*; 2) clefts are more sensitive to context integration than sentences with the adverb. In the AJTs, there were sets of stimuli in four conditions: exhaustive/non-exhaustive ‘only’-sentences (‘*Yesterday, only Anna brewed coffee and pastries too*’/‘*Yesterday only Anna brewed coffee and Olivia did too*’); and exhaustive/non-exhaustive clefts *(‘Yesterday it was Anna who brewed coffee and pastries too’/ ‘Yesterday it was Anna who brewed coffee and Olivia did too’*). In the eye-tracking experiment, we included context-sentences, displayed on previous screens (*‘Anna and Olivia run a busy bakery in London’*). Results show that: 1) ‘only’-sentences were less accepted and elicited more regression-based measures than clefts; 2) non-exhaustivity decreased the acceptability of all sentences, and reduced the reading times for adverb-sentences. Our findings corroborate the idea that exhaustivity relates to the truth values in ‘only’-sentences (Rooth, 1992; Kriffa, 1999), whereas to pragmatic in clefts (Drenhaus; Zimmerman; Vasishth, 2010; Onea, 2019). We argue that exhaustivity in clefts engage *prediction-by-production* (Pickering; Gambi, 2018), in which comprehenders simulate the speaker’s communicative intention (Destruel *et al*., 2015; Destruel; Tunnicliff, 2020). However, exhaustivity in ‘only’-sentences is part of the core semantic content, unleashing the fast, automatic process of *prediction-by-association* based on associations between words or structures (Pickering; Gambi, 2018).

Keywords: Focus particles. Clefts. Exhaustivity. Prediction.

RESUMO

Este estudo investiga dois tipos de estruturas: sentenças com o advérbio ‘só’/‘*only*’ e clivadas totais*/it-clefts*, em inglês e em português brasileiro (PB). Há evidências de que esses marcadores de foco (Kiss, 1998; Paterson Liversedge; Underwood, 1999; Liversedge *et al.*, 2002; Zimmerman, 2004; Filik; Paterson; Liversedge, 2009; Carlson, 2015) convergem exaustividade por motivos distintos (Horn, 1981; Rooth, 1992; Krifka, 1999; Drenhaus; Zimmerman; Vasishth, 2011; Carlson, 2014; Lignani; Fonseca, 2023). ‘Exaustividade’ é o efeito semântico através do qual somente o elemento focalizado de uma sentença a torna verdadeira, excluindo todas as outras alternativas contextualmente possíveis (Rooth, 1992; Krifka, 1999). Desenvolvemos dois tipos de experimentos para investigar violações de exaustividade: a) três tarefas de julgamento de aceitabilidade – AJT (em inglês, tarefas 1 e 2; em PB, tarefa 3); e b) duas tarefas com rastreamento ocular (em inglês e em PB). Nossas hipóteses eram: 1) violações de exaustividade em sentenças-‘só’ são menos aceitáveis do que em clivadas totais; 2) clivadas totais são mais sensíveis à integração de contexto do que sentenças ‘só’. Para as AJTs, utilizamos conjuntos de estímulos em quatro condições diferentes: sentenças com ‘só' exaustivas/não-exaustivas (‘*Ontem, só a Anna preparou o café e também o chá*’/‘*Ontem, só a Anna preparou o café e também a Olívia*’); e clivadas exaustivas/não-exaustivas (‘*Ontem, foi a Anna quem preparou o café e também o chá*’/‘*Ontem, foi a Anna quem preparou o café e também a Olívia*’). No experimento de rastreamento ocular, incluímos condições neutras exaustivas/não-exaustivas (idênticas às sentenças com marcadores de foco, à exceção destes) e frases de contexto, exibidas em uma tela anterior à que continha o item experimental (‘*Anna e Olivia têm uma padaria movimentada em Londres*’). Observou-se que: 1) sentenças com ‘só’ foram menos aceitas e provocaram mais regressões do que clivadas; 2) a não-exaustividade diminuiu a aceitabilidade dos itens em todas as condições e o tempo de leitura de sentenças com o advérbio. Esses resultados corroboram a ideia de que a exaustividade está relacionada às condições de verdade em sentenças com ‘só’ (Rooth, 1992; Krifka, 1999) e a implicaturas pragmáticas em clivadas (Drenhaus, Zimmerman e Vasishth, 2010; Onea, 2019). Argumentamos que clivadas podem desencadear *predição-por-produção*, que implica simulação da intenção comunicativa, enquanto que sentenças exclusivas são processadas segundo a *predição-por-associação*, baseada em associações entre palavras e/ou estruturas (Destruel *et al*., 2015; Pickering; Gambi, 2018; Destruel; Tunnicliff, 2020).

Palavras-chave: Advérbios focalizadores. Clivagem. Exaustividade. Predição.

SUMMARY

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This thesis comprehends topics in the area of Experimental Psycholinguistics (Traxler, 2013) and Experimental Semantics and Pragmatics (Cummins; Katsos, 2019), aiming it-clefts/*clivadas totais* and sentences with the focus particle 'only*'/'só*', in English and in Brazilian Portuguese (BP). The broad objective of this research is to investigate how people process exhaustivity in these kinds of structures.

According to Rooth's (1992) alternative semantics, focus introduces a set of contextually relevant alternatives, which are licensed by a covert semantic operator. A focus marker, such as 'only' or clefts, associates with this operator and identifies a specific (focused) constituent for which the sentence is true. The focused constituent contrasts with a group of contextually presupposed elements, and that is the idea of the contrastive focus, which originates the semantic exhaustivity (Carlson, 2014).

Sentences with 'only' are exhaustive due to the lexical insertion of the adverb in the syntactic structure, which overtly marks an element and supposedly affects the truth-conditional values of the sentence (Rooth, 1992; Krifka, 1999). The idea of exclusivity from 'only' makes exhaustivity to be grammatically encoded and theoretically obligatory for sentences with such adverbs (König, 1991). Differently, in clefts, exhaustivity is typically understood to be derived from a pragmatic effect, conversational implicatures, rather than a truth-conditional effect (Horn, 1981; Kiss, 1998; Krifka, 1999; Pollard; Yasavul, 2016; Onea, 2019). Therefore, although both clefts and 'only'-sentences are capable of conveying exhaustivity, there is evidence to believe that they differ in whether it is semantically encoded or pragmatically inferred (Krifka, 1999; Drenhaus; Zimmerman; Vasishth, 2010; Destruel *et al.*, 2015; Lignani, 2023).

It is important to consider that exhaustivity processing happens through predictions. According to Pickering and Gambi (2018), linguistic prediction corresponds to a pre-activation of syntactic, semantic and/or phonological linguistic information before the comprehensioners perceive the linguistic stimuli that carries such information. Both 'only'-sentences and clefts trigger expectations of exhaustivity during processing, but the evidence that they convey exhaustivity for different reasons might lead one to believe that they rely on different predictive mechanisms to do so.

In ‘only’-sentences, exhaustivity is lexically encoded and immediately accessible, which leads comprehenders to generate strong, early predictions about the exclusivity of the focused element. This might be related to what Pickering and Gambi (2018) define as *prediction-by-association*. In contrast, if clefts truly convey exhaustivity through pragmatic implicature, they require comprehenders to infer the speaker’s communicative intent and evaluate the context of the sentence, often engaging *prediction-by-production* (Pickering; Gambi, 2018). However, in both cases exhaustivity is shaped by linguistic cues that guide predictive processing, and in certain cases these cues might even include the contextual information (Krifka, 2008; Onea, 2019).

There are studies in Psycholinguistics that provide evidence to online and *in-locus* focus processing (Paterson Liversedge; Underwood, 1999; Liversedge *et al*., 2007; Filik *et al*., 2009; Drenhaus; Zimmerman; Vasishth, 2010; Carlson, 2014; 2015; Lignani; Fonseca, 2023). Therefore, if adverbs and clefts associate in different ways with the semantic operator of exhaustivity, focus in each type of sentence might be predicted in different ways and/or in different locations.

Moreover, there are several studies comparing exclusive adverbs, such as ‘only’ and ‘só’ in different languages (König, 1991; Cinque, 1999; Kiss, 1998; Cesare, 2010; Quarezemin; Tescari Neto, 2015), as well as clefts (Hedberg, 2000; Alves *et al*., 2015; Destruel *et al*., 2015; De Veaugh-Geiss *et al*., 2018). These cross-linguistic studies suggest that adverbs and clefts operate in similar ways across languages, which is why we developed this research in English and in BP. We intended to look for evidence of similarities and/or differences between these two languages.

The specific objectives for this research were: i) to contrast ‘only’-sentences and clefts, regarding (non-)exhaustivity; ii) to contrast data obtained in English and in BP; iii) to investigate how previous context influence the interpretation of exhaustivity in those syntactic structures; iv) to develop a model of exhaustivity processing that describes these two syntactic-semantic strategies model.

In order to achieve the specific objectives, we developed three experiments in English and two in BP, all described in detail in Section 4. Two of the English experiments and the BP one were the so-called ‘acceptability judgment tasks’ (AJTs). These were Likert scale tests (Likert, 1932) in which participants should rate ‘only’-sentences and clefts according to how acceptable they judged them to be. Stimuli contained decontextualized sentences that either

fulfilled the so-predicted exhaustivity or presented a violation to it, which provoked a disruption in people's expectations.

The other kind of experiment conducted was an eye-tracking task, in which, again, participants would read 'only'-sentences and clefts that fulfilled or violated exhaustivity, but in this experiment all stimuli were presented after context-sentences.

In this document, the following sections describe the theoretical background, the method and the results of this project. The sections of this paper are organised as following: **1 INTRODUCTION**; **2 SENTENCE PROCESSING**; **3 INFORMATION STRUCTURE**; **4 METHODOLOGY**; **5 METHODOLOGICAL DISCUSSIONS AND FURTHER DIRECTIONS**; **6 LAST CONSIDERATIONS**; **REFERENCES**.

2. SENTENCE PROCESSING

This research is in the interface of the areas of Experimental Psycholinguistics – specifically, *sentence processing* – and Experimental Semantics and Pragmatics.

Psycholinguistics investigates how language is processed in real time, focusing on the cognitive mechanisms that underlie comprehension and production (Traxler, 2013). Rather than examining only the representations encoded in linguistic systems, this field concerns how linguistic information is incrementally integrated, how expectations are generated, and how interpretations are updated during processing.

Sentence processing refers to the set of cognitive operations involved in the real-time comprehension of sentences. These operations include lexical access, syntactic parsing (or simply *parsing*), semantic integration, memory retrieval, inferencing, prediction, and pragmatic interpretation (Traxler, 2013). P*arsing* is the process through which comprehenders incrementally assign structure to linguistic *input* as words are read and/or heard, thus it is responsible for online construction and analysis of sentence structure during language comprehension (Wilson; Keil, 2001). Therefore, sentence interpretation involves the *parsing* and multiple cognitive processes, such as pragmatic integration.

A central question in Psycholinguistics concerns the extent to which sentence comprehension is immediate and fully specified, or instead guided by heuristic and predictive processes. Some models propose that comprehenders build detailed representations as soon as possible (Frazier, 2002), whereas others suggest that they rely on partial and underspecified interpretations that are revised when necessary (Ferreira; Bailey; Ferraro, 2002). Predictive approaches argue that readers continuously generate expectations about upcoming input by using linguistic cues and contextual information (Pickering; Gambi, 2018).

In recent decades, semantics and pragmatics have shifted from relying primarily on introspective judgments to incorporating experimental methods. Within this context, Experimental Semantics and Pragmatics has emerged as a subfield of experimental linguistics that aims to provide empirical evidence capable of informing theoretical debates about meaning (Cummins; Katsos, 2019). This field may seem closely related to Psycholinguistics, since both address questions concerning how sentence meaning is computed. However, Psycholinguistics primarily aims to model the cognitive mechanisms underlying language processing (Traxler, 2013), whereas Experimental Semantics and Pragmatics uses empirical methods to test theoretical claims about meaning.

The present study lies at the intersection of these fields, as it combines psycholinguistic methods – particularly, eye-tracking measures of online sentence comprehension – with questions concerning semantic and pragmatic computation, namely how the effect of exhaustivity is processed in focus constructions. Therefore, the present thesis contributes to both areas by investigating how exhaustivity is processed across focus constructions and languages while employing concepts and methods widely explored in Psycholinguistics to test semantic and pragmatic theories.

The next sections explore some of the key-concepts that matter for this research, such as prediction, syntactic reanalysis and integration.

2.1. Reading Tasks

Sentence comprehension unfolds incrementally, as readers[1] continuously interpret incoming linguistic information in real time. During this process, lexical, syntactic, semantic, and discourse representations must remain temporarily available so that new information can be integrated into a coherent interpretation (Just; Carpenter, 1992; Caplan; Waters, 1999). In this sense, memory-related mechanisms[2] support the maintenance and retrieval of linguistic information throughout sentence processing. Moreover, predictive and associative mechanisms contribute to the activation of semantically and/or syntactically related representations, as well as the management of the expectations created by the reader.

Reading tasks provide important evidence for investigating comprehension. In Psycholinguistics, such tasks can be divided into *offline* or *online* methodologies, depending on whether they investigate the final interpretation of a sentence or the implicit processing operations that occur during comprehension (Pliatsikas; Marinis, 2020).

Offline tasks examine linguistic interpretation after comprehension has already taken place. In these tasks, participants engage in reflective decision-making tasks, making judgments, providing answers, after having read the entire sentence or discourse. Likert tasks (Likert, 1932), comprehension questionnaires, paraphrase selection, and truth-value judgment

[1] We will use the term ‘reader’ along this text as a synonym for ‘comprehender’. Although comprehension is a process that might involve acoustic stimuli too - in the case of inputs that are heard - the methods of this research comprise reading tasks.

[2] *Working memory* is a cognitive system responsible to temporarily store and manipulate information required for complex cognitive tasks such language comprehension. According to Baddeley (2000), working memory is composed of multiple components, including a central executive that controls attention and two storage subsystems: the *phonological loop* (for verbal/auditory information) and the *visuospatial sketchpad* (for visual/spatial information).

tasks are common examples of offline methodologies. Because participants access the final outcome of comprehension, offline tasks are particularly useful for investigating interpretive preferences, grammaticality judgments, semantic perceptions, and conscious metalinguistic evaluation. However, such tasks give limited information about the intermediate cognitive operations involved during processing, such as prediction, temporary ambiguity resolution, reanalysis, or incremental integration.

On the other hand, online tasks investigate comprehension as it unfolds in real time. These methodologies aim to capture the moment-by-moment cognitive processes involved in sentence interpretation, allowing researchers to examine how linguistic information is incrementally processed during reading or listening. *Self-paced reading* and *eye-tracking* are among the most widely used online methodologies in sentence processing research. Self-paced reading tasks typically measure reading times across sentence regions (Just; Carpenter; Woolley, 1992), whereas eye-tracking provides more fine-grained evidence through measures such as fixation durations, regressions, gaze patterns, and regression-path durations (Rayner, 1998). Because online methodologies access intermediate stages of processing, they are particularly useful for investigating prediction, *parsing*, semantic integration, processing difficulty, expectation violation, and reanalysis effects.

Importantly, online and offline tasks often provide complementary evidence. While offline tasks reveal how comprehenders ultimately interpret linguistic input, online methodologies make it possible to investigate the cognitive processes that lead to such interpretations. This distinction is especially relevant for studies such as the present one, which includes focus, exhaustivity and pragmatic inference: comprehenders may show temporary processing costs or intermediate interpretations that are not necessarily reflected in final offline judgments.

Eye-tracking methodologies provide more fine-grained evidence about online comprehension by examining measures such as fixation durations, regressions, regression-path duration, and gaze patterns during reading (Rayner, 1998). This task has become one of the main methodologies for investigating sentence processing in real time.

There is robust evidence to the correlation between complex reading demands with longer fixation durations and increased regressions, particularly when processing complex embeddings or resolving antecedents in long-distance dependencies (Van Dyke; McElree, 2006; Traxler; 2013). Moreover, retrieval processes interact with attentional control

mechanisms to manage interference effects from similar distractors during *parsing* (Dehaene, 2020). This control is critical when multiple plausible interpretations compete, as in temporary ambiguity resolution or other cases that demand more cognitive involvement, such as *VP-ellipsis* and *gapping*, two structures that were used for the experimental items in this research, as it will be discussed in Section 4.

Verb-phrase ellipsis (*VP-ellipsis*) and *gapping* involve the omission of linguistic material that must be mentally reconstructed during comprehension to form an efficient interpretation. In VP-ellipsis (*e.g.*, 'John bought apples, and Mary did too'), the verb-phrase of the first clause ('bought apples') must be retrieved and integrated into the second clause. This retrieval process represents a cost to working memory, especially in cases involving distant or complex antecedents (Frazier; Clifton, 2001; Martin; McElree, 2009).

Similarly, in *gapping* structures (*e.g.*, 'John bought apples, and also Mary'), the verb-phrase ('bought apples') must be recovered and assigned correctly to the second conjunct, which requires syntactic and semantic reactivation of the elided content (Carlson, 2002). Studies using self-paced reading and eye-tracking have shown that increased distance between the ellipsis site and its antecedent leads to longer reading times and increased processing cost, further supporting the involvement of working memory resources (Martin; McElree, 2009). These findings suggest that efficient ellipsis resolution depends not only on syntactic structure but also on the capacity to maintain and retrieve linguistic representations.

As individuals read, they continuously integrate incoming linguistic *input* with contextual cues and prior knowledge, constructing a dynamic mental representation of the discourse. This continuous process involves both prediction and integration, processes of higher-order that will be later discussed, and fundamentally supported by interactions between *bottom-up* and *top-down* processes.

Bottom-up processing involves the incremental interpretation of linguistic input based solely on the information available in the signal – lexical, morphological, syntactic, and prosodic features. This type of processing is essential for constructing a local syntactic and semantic representation of the utterance. For instance, readers use syntactic cues to assign grammatical roles and interpret sentence structure as it unfolds (Frazier; Fodor, 1978). In contrast, *top-down* processing draws on stored knowledge – including semantic associations, world knowledge, discourse context, and expectations about upcoming input – to guide and constrain interpretation. For example, when reading a sentence like 'The teacher reached for

the…', readers are more likely to expect a word like 'pen' or 'paper' than an unrelated object, due to their knowledge about typical events in schools. These *top-down* inferences can bias lexical access and syntactic parsing even before *bottom-up* information is fully processed (Hersch; Andrews, 2012).

Since this research uses reading tasks to investigate the processing of exhaustivity in focus constructions, understanding how reading and comprehension relates to concepts such as prediction, integration, inferencing and reanalysis is paramount. The following sections address each one of these processes.

2.2. Garden-Path Effects and Syntactic Reanalysis

A central question in sentence processing research concerns how comprehenders resolve temporary ambiguities perceived during real-time comprehension. One of the most influential frameworks addressing this issue is the *Garden-Path Theory – GPT* (Frazier; Fodor, 1978; Frazier, 1999).

This theory argues that the *parser* initially takes into account only the syntactic information of a sentence to construct its interpretation. GPT argues that the comprehension process is formed by two stages, once the *input* is perceived: i) *Preliminary Phrase Packager* (*PPP*), when the *parser* determines the category of the lexical items and merge them into phrases; and ii) *Sentence Structure Supervisor* (*SSS*), whereby the phrases are combined into a whole syntactic structure (clause). GPT's process is based in principles of processing economy (*Minimal Attachment*[3] and *Late Closure*[4]). According to Frazier (1979), the fact that structures are initially processed with minimal effort explains why the *parser* prioritises information of syntactic nature –supposedly, easier – even though this means potentially incorrect *parsing* of a sentence. Whether the *parser* fails to build a whole sentence that makes sense (during *SSS*), that means that the first processing was incorrect, and that it has to be revised, which represents extra difficulties for this whole process (Frazier; Rayner, 1982; Frazier, 1987).

Garden-path sentences typically involve syntactic ambiguities that compel the processor to commit to one structural interpretation rather than another, only to later

[3] "Minimal Attachment: attach incoming material into the phrase marker being constructed using the fewest nodes consistent with the well-formedness rules of the language" (Frazier, 1979, p. 33).

[4] "Late Closure: when possible, attach incoming material into the clause or phrase currently being parsed" (Frazier, 1979, p. 24).

encounter disambiguating material that needs syntactic reanalysis. Classic examples include reduced relative clauses and noun-phrase coordination ambiguities, such as (1):

1) a. "The horse raced past the barn fell" (Frazier; Rayner, 1982, p. 179).

b. "While Anna dressed the baby played in the crib" (Frazier, 1987).

These sentences exemplify how comprehenders initially misparse the structure based on economy principles and subsequently engage in complex reanalysis to recover the intended meaning (Frazier, 1987; Clifton *et al*., 1991).

Behavioral evidence from eye-tracking (Rayner *et al*., 1983), self-paced reading (Just *et al*., 1996), and *event-related potential*[5] (ERP) studies (Osterhout; Holcomb, 1992) consistently demonstrates elevated processing costs in regions where reanalysis is triggered. These costs manifest as increased fixation durations, longer regression paths, and P600 components, all signaling syntactic reprocessing. Although originally conceived to explain syntactic misanalysis, the garden-path model has since been extended to other domains, including semantic, pragmatic, and informational levels of representation (Christianson *et al*., 2001; Patson; Ferreira, 2009; Kuperberg; Jeager, 2016). For instance, comprehenders can also be misled by strong contextual cues or prosodic patterns that suggest an interpretation, which is later contradicted.

The experiments reported in this thesis project provide novel evidence that garden-path effects can occur at the level of pragmatic inferences, particularly those involving focus-sensitive constructions (such as clefts and 'only'-sentences). In these cases, the processor constructs an initial representation based on a default assumption of exhaustivity, which is later contradicted by an additive continuation (*e.g., 'and Mary did too'*). Consider (2):

2) 'Only Daisy baked a cake and Mary did too*'*.

---

[5] *Event-related potentials* (ERPs) are electrical patterns elicited in the brain during the processing of certain types of linguistic stimuli (Kutas; Van Petten, 2006). They are 'event–related' because some of them identify as related to certain types of linguistic stimuli, occurring only in situations where these stimuli are present. These electrical patterns provide information about: i) latency peaks, that is, the time elapsed between a linguistic stimulus and the peak of the magnetic wave; ii) polarity (N/negative or P/positive peaks); and iii) wave amplitude (Kutas; Van Petten, 2006). Identification and observation of ERPs depend on studies that use electroencephalography (EEG).

In (2), by the time that a comprehender reads the focus marker '*only*', the *parser* predicts an exclusive focus interpretation, which means that Daisy was the sole person who baked a cake. The subsequent phrase, '*and Mary did too*', violates this assumption and forces the *parser* to revise the interpretation. The *parser* did so by accessing pragmatic information to complement what is being missed by syntax. Although no ambiguity is present, there is a mismatch between what the *parser* predicted (expectation) and the encountered *input* could, and this could trigger a garden-path-like effect, forcing syntactic reanalysis. Such effects are crucial to understand syntactic reanalysis and semantic cancellation overall.

2.3. Good-Enough Approach

The *Good-Enough Approach* – GE (Ferreira *et al*., 2002; Ferreira; Patson, 2007) – proposes that language comprehension did not always entail the construction of fully specified or syntactically complete representations – differently from the GPT argues. Instead, comprehenders often settle for partial or heuristic-based interpretations that are good enough for the task. This view stands in contrast to models assuming that sentence processing routinely aims for exhaustive and accurate parses, as the GPT, for instance, suggests that any inaccuracy would necessarily unleash a process of reanalysis.

This approach emerged in response to evidence showing that comprehenders sometimes fail to detect implausibility or inconsistencies in syntactic and/or semantic structure, particularly when these do not interfere in the overall comprehension. In this framework, processing is fast, frugal, and goal-oriented, favoring efficient communication over formal correctness (Christianson *et al*., 2001; Ferreira; Bailey; Ferraro, 2002).

Empirical support for the GE comes from a variety of paradigms, including eye-tracking, self-paced reading, and recall tasks. One of the most cited examples involves misinterpretation of garden-path sentences like (1b) "While Anna dressed the baby played in the crib*"* (Frazier, 1987). Participants of experiments with items like these often report an incorrect interpretation in which Anna is dressing the baby even after reading the entire sentence. This suggests incomplete or unsuccessful reanalysis (Christianson *et al*., 2001). These findings challenge the assumption that reanalysis always succeeds in producing a fully revised mental representation.

ERPs studies have shown also that comprehenders sometimes show neural signatures consistent with partial interpretation even when the *input* is grammatically anomalous,

pointing to a dissociation between initial parse commitment and post hoc revision (Van Herten; Kolk; Chwilla, 2006).

The GE model rests on two core assumptions: i) *Satisficing Parsing;* and ii) *Heuristic Use*. The first one suggests that rather than computing a complete parse, the processor stops once they achieve a coherent or plausible interpretation, even though it is incorrect. According to *Heuristic Use*, comprehenders rely on lexical, semantic, and pragmatic cues – such as subject-verb proximity, thematic role prototyping, and real-world knowledge – to guide interpretation, especially under time pressure or cognitive load (Ferreira; Lowder, 2016). Therefore, errors in comprehension are not mere failures, but reflections of strategic adaptation in resource-constrained environments.

Despite their differences, the Garden-Path Theory and the Good-Enough Approach share a core assumption: sentence processing involves rapid heuristic-based commitments, which may later require revision. Both frameworks highlight the fallibility of real-time comprehension and emphasize the interplay between incremental interpretation and reanalysis mechanisms. These insights converge with integration-based and predictive accounts, which propose that comprehenders continuously generate expectations and integrate diverse sources of information to convey meaning. In the next sections, we will discuss linguistic prediction and integration.

2.4. Predictive Processing and Integration

The notion that language comprehension involves prediction has gained substantial empirical and theoretical support in the past two decades. Predictive processing theories argue that comprehenders do not passively wait for linguistic input to accumulate but instead actively anticipate upcoming information at multiple levels: phonological, lexical, syntactic, and discourse-pragmatic. This approach marks a shift from strictly *bottom-up*, *input*-driven accounts of language processing to more interactive, expectation-driven models, where *top-down* information such as context, world knowledge, and prior linguistic cues modulates comprehension in real time (Altmann; Kamide, 1999; Kutas; Van Petten, 2006; Pickering; Gambi, 2018).

Predictive processing emerges as a natural consequence of *top-down* integration. Rather than passively receiving linguistic *input*, comprehenders actively create expectations about upcoming material. These predictions can target upcoming words (Altmann; Kamide,

1999), syntactic structures (Levy, 2013), or even discourse moves (Nieuwland; Van Berkum, 2006). Importantly, prediction is not uniform across individuals or contexts. It interacts with memory availability, lexical frequency, syntactic complexity, and individual cognitive capacities (Staub, 2015). Furthermore, predictions may be influenced by prior discourse and long-term associations (Van Gompel; Majid, 2004).

In Psycholinguistics, behavioral and neurophysiological studies have provided compelling evidence that readers and listeners generate linguistic predictions. In visual world paradigms (Altmann; Kamide, 1999; Maia, 2008), participants fixate on semantically plausible objects in a scene even before these objects are mentioned, suggesting that verb semantics guide expectations about upcoming noun phrases. ERPs' studies also support this claim: the N400 component, typically associated with semantic integration difficulty, is attenuated when upcoming words are more predictable in context (Kutas; Hillyard, 1984; Delong; Urbach; Kutas; Van Petten, 2006). Similarly, eye-tracking studies in reading have shown shorter fixation durations and reduced regressions on highly predictable words (Rayner *et al*., 2004; Clifton, Staub; Rayner, 2007), indicating that prediction facilitates the ease of lexical access and syntactic parsing.

Predictive processing is tightly linked to incremental integration, the process by which incoming *input* is continuously incorporated into a developing mental representation of the sentence. Differently, integration operates at multiple levels and requires the comprehender to reconcile syntactic, semantic, and discourse-level constraints. The integration of new material is shaped not only by the *input* itself, but also by prior expectations. For instance, when predictions are disconfirmed – such as when an unexpected noun appears after a restrictive verb – comprehension slows down and often triggers reanalysis or repair processes (Kuperberg; Jaeger, 2016). Integration hence mediates the balance between what is anticipated and what is encountered.

While traditional models such as the Garden-Path focus on syntactic misanalysis and reanalysis, integration-based models emphasize how multiple information streams are being accessed and coordinated to generate the most coherent interpretation at each moment. This view aligns with constraint-based theories of sentence processing (MacDonald *et al*., 1994), which hold that the *parser* uses all available probabilistic cues to guide comprehension, including lexical frequency, plausibility, and prosody.

Predictive mechanisms are particularly relevant in this research, as our scope is in the processing of focus-sensitive constructions, such as clefts and '*only*'-sentences. These constructions often give rise to exhaustive inferences, interpretations that not only assert what is true but also imply that other alternatives are false. Because such inferences depend on the informational structure of a sentence, they offer a valuable testing ground for understanding the timing and depth of predictive mechanisms.

For instance, studies suggest that listeners may anticipate the presence of contrastive alternatives upon encountering focus cues, which then shapes the interpretation of the focused element (Fraundorf; Watson; Benjamin, 2010; DeVeaugh-Geiss *et al*., 2018). In this sense, prediction is not merely about upcoming words or structures, but also about upcoming semantic and pragmatic commitments. This connects with the GE approach, in which comprehenders sometimes generate shallow interpretations, based in plausible expectations, and engage in deeper analysis only when these expectations are violated or underspecified. As we consider violations of exhaustivity, a semantic operator present in every sentence, the way that focus markers associated with this operator might reflect in how people make predictions about it.

It is now important to make a distinction: although dynamically related, integration and prediction are conceptually distinct. Predictions involve anticipation, pre-activating potential input based on the current context, whereas integration refers to the incorporation of actual input into the developing mental representation (Pickering; Gambi, 2018). Successful language comprehension requires a fine-tuned balance between the two: over-reliance on anticipation may lead to errors (as in garden-path sentences), while purely reactive integration can result in delayed or effortful comprehension. From a processing perspective, anticipatory mechanisms may reduce the computational load at the moment of input (in cases when the anticipation was successful). On the other hand, integrative mechanisms serve as a corrective and updating system.

The extent and nature of predictive processing may vary across languages and constructions. For example, word order flexibility, prosodic structure, and information structure strategies affect how much a language invites or constrains prediction. Studies on cleft constructions in different languages (Destruel *et al*., 2015; Greif; Skopeteas, 2021) have shown that informational focus provided by clefting interacts with context, shaping both expectations and processing effort. This cross-linguistic variation underscores the need to

examine language-specific cues and their interaction with universal cognitive principles in predictive comprehension.

2.5. Prediction Theory

Pickering and Gambi (2018) propose that language comprehension involves prediction, meaning that listeners and readers actively anticipate upcoming linguistic *input*. This predictive mechanism allows for faster and more efficient processing, as the system can pre-activate possible continuations of the sentence before they occur. In their model, prediction is central to language understanding and is informed by both syntactic and semantic cues available in context. There is also a distinction of two types of prediction: an optional one, *prediction-by-production*, and an ubiquitous one, *prediction-by- association*.

Prediction-by-production would occur based on a system of imitation of the linguistic form of a produced utterance by the comprehender. This would lead to the construction of the representation of the speaker's underlying communicative intention. The comprehension goes through this process of 'covert imitation' (Pickering; Gambi, 2018): the production representations correspond to the representations constructed by the comprehension system. From this, the comprehenders consider contextual elements and derive the intention of the speaker's utterance, coming up with an interpretation. This type of linguistic prediction, although optional, tends to be accurate and effective, as it converges mechanisms used in the production of the language itself (Pickering; Gambi, 2018).

Conversely, in the *prediction-by-association* process, prediction occurs through the network activation of correlated syntactic representations, analogously to *priming effects* (syntactic, semantic and/or phonological). This prediction system is not compatible with prediction-by-production, since the latter considers the conversational context and network activation can automatically activate information that is semantically inappropriate to the context, but syntactically and/or phonologically similar to the *input* previously processed. Considering the possibility of semantic inadequacy of the predicted constituents and the ubiquitous nature of spread activation, *prediction-by-association* is automatic (involuntary), but less effective, that is, with a high error rate.

Although the Prediction Theory (Pickering; Gambi, 2018) does not focus specifically on clefts or sentences with focus adverbs, such as 'only', their model provides a useful framework for understanding how comprehenders might anticipate the presence of

exhaustivity. Structures like ‘only’-sentences and clefts often imply a contrast with alternatives, thus readers might predict that a statement is exhaustive. For instance, after reading a context sentence introducing multiple agents, a reader might expect a contrastive structure and interpret a cleft or an ‘only’-sentence accordingly. This aligns with the idea that semantic properties like exhaustivity can be anticipated, as part of broader meaning-level predictions.

A key distinction in Pickering and Gambi’s model (2018) is between prediction and integration. The authors strengthen the differences that we discussed in the previous section, arguing that ‘prediction’ occurs before the linguistic *input* is encountered, and reflects the anticipatory nature of comprehension. Integration, in contrast, happens after the *input* is perceived and involves fitting the new information into the mental representation of the discourse. While prediction can reduce processing load when expectations are met, it can also represent extra costs for the *parser* when it anticipates information that is not presented. Likewise, integration essentially represents more effort than prediction, but it can actually be very helpful and reduce difficulties of an interpretation after reanalysis.

These two processes are therefore both essential and functionally distinct: prediction facilitates comprehension by preparing the system, while integration ensures coherence and interpretive accuracy. In this framework, it is also important to distinguish *prediction-by-production* from integration. *Prediction-by-production* occurs before the linguistic input is encountered: listeners or readers use their internal production system to simulate what a speaker is likely to say next, based on contextual cues and communicative goals. This process involves anticipating specific meanings or structures, such as the use of only to signal exhaustiveness. Integration, by contrast, happens after the *input* has been processed – it refers to how comprehenders update their mental representation of the sentence by incorporating new information. While *prediction-by-production* is forward-looking and intention-based, integration is retrospective and involves coherence-building.

We have discussed some concepts and perspectives regarding sentence processing that are important for this research. In the next Section, we will present the objects of our studies: sentences with focus adverbs and clefts.

3. INFORMATION STRUCTURE

The focus of a sentence carries the notion of figure and background, and it can be manipulated with certain syntactic strategies (Ilari, 2002). According to Kiss (1998), clefting and the use of focus particles (such as adverbs) are among these strategies. Clefts allow a change in the canonic order of a sentence, therefore focusing on a constituent. The focused constituent is then consequently highlighted exhaustively (Carlson, 2015).

The truth values of a sentence concern its informational structure, organized based on logical relationships. In this way, each lexical constituent of a sentence has a specific semantic value and the sentence has the semantic value of a proposition (Rooth, 1992). Semantic or pragmatic rules can use the semantic value of focus, determined in the syntactic structure of a sentence. For focus adverbs, there is robust evidence that the semantic effect of Exhaustivity caused by them relates to the logical structure of the sentence (König, 1991; Rooth, 1992). Clefts, however, do not seem to have the exhaustivity relationship built in the same way.

According to Krifka (2008), the so-called *semantic use of focus* relates to the factual information, which is the content of the sentence itself. Focus-sensitive operators, such as 'only', inherently encode an exhaustive interpretation: they assert that the focused constituent is the one exclusive true alternative among a relevant set of alternatives. This interpretation is encoded in the lexical-semantic level, when the truth-conditions operate to form a grammatical sentence. The presupposed elements are those contextually possible (Kiss, 1998).

For example, consider the it-cleft (3) and the 'only'-sentence (4):

3) "It was MARY who brought the book[6]" (Alves *et al.,* 2015).

4) 'Only MARY brought the book'[7]

The constituent 'Mary' is exhaustively marked, presenting itself as a contrast to other possible constituents, such as 'Anna' or 'Linda', in a context where three students (Mary, Anna and Linda) go to school together. Similarly, in (4) the focus particle 'only' is a similar strategy to clefts, converging exhaustivity to 'Mary', contrastively to ' Anna' and 'Linda'.

Differently, the *pragmatic use of focus* would rely on contextual assumptions, because it associates with the communicative goals (Krifka, 2008). This type of focus marks the new part of the information, operating in the discourse, without changing the content of the

[6] Example translated from Alves *et al*. (2015). The original, in BP ('*Foi a Maria quem trouxe o livro*') is called a 'clivada total' sentence, which has a syntactic/semantic effect analogous to the English it-cleft.

[7] In BP, "Só a Mary trouxe o livro".

proposition. That is why, with supportive context, the *pragmatic use of focus* is also capable of confirming and correcting information (Krifka, 2008).

The author proposes a new type of focus, other than *semantic* or *pragmatic*, which he calls *exhaustive focus*. It refers to the specific interpretation of the alternative's contribution: "the focus denotation is the logically strongest that did so" (Krifka, 2008, p.17). In other words, *exhaustive focus* is a consequence of the interpretation of exhaustivity, which might be semantic (structural) or pragmatic (implicature). The clefts (in English) are Krifka's (2008) example of *exhaustive focus*, because they do not necessarily convey exhaustivity – as focus particles, such as 'only' – but are frequently processed as exhaustive sentences, probably because of their conversational implicatures (Horn, 1981). Overall, even if for different reasons, both the clefts and focus adverbs convey exhaustivity.

According to the literature in Psycholinguistics, both clefts and focus change the informational structure of the sentences (Rooth, 1992), which is considered during sentence processing (Paterson; Liversedge; Underwood, 1999; Liversedge, Paterson; Clayes, 2002; Paterson *et al.*, 2007; Filik, Paterson; Liversedge, 2009; Drenhaus, Zimmermann and Vasishth, 2010; Carlson, 2015). These changes convey exhaustivity and are perceived by readers/listeners, which influence the linguistic prediction of unaccessed sentential constituents (Carlson, 2014; Ryskin, Levy; Fedorenko, 2020; Lignani; Fonseca, 2023).

3.1. Focus adverbs 'only' and '*só*'

Focus particles, such as '*só*'/'only', highlight a sentential constituent, associating with it. '*Só*'/'*only* ' works as an operator that translates a scalar and exclusiveness relationship, since it marks the constituent that on its scope as the only satisfactory (exclusive) option, within a range of contextually accepted possibilities (scale) for which the sentence is true (Kiss, 1998). This constituent becomes one that, within a specific context, is marked exhaustively and has the properties to make the sentence true (Carlson, 2015).

De Cesare (2010), following König (1993), highlights that focus adverbs function at the interface of syntax, semantics, and prosody. For '*only*' and its counterparts, it is common that the focus is marked syntactically and prosodically, typically through pitch accents. This prosodic marking is crucial for disambiguating the focus scope and associating with the adverb. In BP, prosody also plays a decisive role in identifying the focused constituent,

especially when syntactic cues are ambiguous. Studies of spoken language show that pitch, duration, and intensity signal focal prominence, which is essential for correctly interpreting '*só*' in context.

In *alternative semantics* (Rooth, 1992), 'only' is analysed as an operator that associates with focus and triggers an exhaustive inference. It takes a focused constituent as its syntactic argument and presupposes a set of alternatives, asserting that the proposition is true only for the focused element. For example, consider (5):

5) 'Only Gabriel came to the party'.

The statement is that Gabriel came to the party, and the implication is that no one else did so. The position of 'only' affects its scope and interpretation (e.g., '*Only John introduced Mary to Sue'* vs. '*John only introduced Mary to Sue'*), and focus placement interacts tightly with prosodic prominence and syntactic structure (Carlson, 2014; Lignani; Fonseca, 2023). In BP, the focus adverb '*só*' typically precedes the constituent it modifies (*e.g.,'Só a Maria veio'*/ '*Only Maria came'*), yielding an interpretation equivalent to English 'only'. However, '*só*' placement can be more flexible, especially in spoken language ('*Maria só veio'*), and its interpretation depends on prosody and context (Souza, 2004).

'Only' and 'só' require the comprehensioner to generate a set of relevant alternatives and evaluate which ones are excluded by the speaker's utterance. This process is cognitively costly and draws on working memory and inferencing mechanisms (Baddeley, 2010). The integration of these alternatives depends not only on *bottom-up* cues (e.g., word order, morphology) but also on *top-down* information such as discourse coherence, speaker intent, and world knowledge. This interaction exemplifies the dual role of focus adverbs in syntactic processing and pragmatic inference. The semantics of such focus adverbs has been extensively studied, both theoretically and empirically, with early psycholinguistic research focusing on how readers and listeners interpret the exhaustivity that the adverbs convey. One of the most influential findings is that the processing of 'only'-sentences is incrementally demanding and sensitive to syntactic and prosodic cues.

Paterson, Liversedge and Underwood (1999) – and later Paterson, Liversedge and Rayner (2007) – conducted eye-tracking experiments showing that readers experienced longer fixation times and more regressions when interpreting sentences, especially when the focal element was ambiguous or mismatched with contextual expectations, like (6):

6) ‘Only Jane ate the cake’

These processing difficulties suggest that comprehenders actively compute exhaustivity in real time and that this computation involves integrating the focus operator with both syntactic structure and contextual alternatives.

Filik, Paterson and Liversedge (2009) further demonstrated that ‘only’-sentences are processed more easily when the focus is disambiguated by prosody or punctuation, highlighting the predictive nature of focus assignment during sentence comprehension. Their results suggest that comprehenders make early commitments about where the focal stress should fall and experience reanalysis costs when those predictions are not confirmed. This aligns with findings by Carlson (2014; 2015), who investigated how prosodic prominence guides listeners’ expectations about focus structure and exhaustivity, indicating a close interaction between focus, prosody, and processing dynamics. Clifton *et al*. (2006) also found that discourse context can license or clash with the presuppositions of only, influencing processing load and interpretation outcomes.

In eye-tracking studies, for instance, readers tend to show shorter fixation durations when the continuation confirms the exclusive inference (Fraundorf; Watson; Benjamin, 2010), and increased regressions or longer reading times when the continuation contradicts that inference (Kim; Levy; Patson, 2015). These results suggest that ‘*only*’ triggers immediate semantic predictions, potentially facilitated by focused prosody and contextual cues. Moreover, studies show that the presence of a focus-sensitive particle helps information to be encoded in memory (Spalek; Gotzner; Wartenburger, 2012).

Crosslinguistically, the behavior of focus adverbs remains relatively stable, though subtle differences have been observed. For instance, Dimroth (2004) and Skopeteas and Fanselow (2010) show that German, English, and Romance languages exhibit similar constraints on focus-sensitive elements, even though the syntactic positions and prosodic realizations vary. In BP, ‘só’ mirrors ‘only’ in its semantic contribution but may differ in scope interpretation depending on word order and intonation (Santos, 2002; Reis, 2013).

Recent experimental work has explored how speakers of BP interpret sentences with só, revealing sensitivity to focus structure and contextually relevant alternatives. Biezma and Rawlins (2012) found that participants consistently derive exhaustivity inferences from só-sentences but are also affected by contextual manipulations that highlight or suppress

alternatives. This supports the idea that focus interpretation is deeply intertwined with discourse modeling and predictability.

Although not exclusively about '*só*', Alves *et al*. (2022) investigates the multimodal expression of focus in BP, showing that both production and perception of focus involve not only prosody but also visual and gestural cues. In their experiments, contrastive focus was reliably identified by listeners through auditory and visual channels, indicating that focus marking in BP is a robust multimodal phenomenon. Moreover, Lignani and Fonseca (2023) provide evidence to the fact that focus is processed in-locus and therefore the variation in the position of the focus markers implies prediction of contrastive syntactic structures. This reinforces the idea that focus adverbs like '*só*' operate within a system of linguistic and paralinguistic cues that jointly convey information structure.

Souza (2004) examines a range of focus-sensitive adverbs in BP – including '*só*', '*mesmo*' ('even'), *até* ('even'), and '*também*' ('too') – and argues that these markers systematically contribute to the structuring of information. There seem to be a tendency for '*só*' to appear in pre-focal position and co-occur with prosodic prominence and reordering strategies, such as clefts, to highlight contrastive or restrictive interpretations. This suggests that '*só*' is not merely a syntactic marker but a pragmatic device that interacts with broader discourse goals. However, unlike English '*only*', which has been extensively tested in comprehension studies, BP '*só*' lacks a robust body of experimental research addressing its real-time processing.

The predictability dimension is crucial: listeners and readers continuously anticipate where focal stress is likely to fall and what information is being excluded. This anticipation is guided by structural cues, prosodic contours, and discourse structure. Focus adverbs, by imposing an exhaustivity presupposition, increase the need for predictive mechanisms that resolve ambiguity and confirm or disconfirm contextual expectations. This aligns with broader theories of predictive processing (Kuperberg; Jaeger, 2016; Pickering; Gambi, 2018), in which language comprehension is viewed as a process of generating, updating, and revising probabilistic hypotheses about linguistic input.

Focus adverbs, like 'only' and '*só*' are powerful triggers of exhaustivity and have measurable effects on processing, both in offline interpretation and in real-time comprehension. Their study has revealed the complex interplay between syntax, prosody, semantics, and context, as well as the role of prediction in focus resolution. These findings

provide a rich framework for investigating how linguistic structures guide attention, restrict alternatives, and shape sentence meaning across languages.

Drenhaus, Zimmermann and Vasishth (2010) shows that, although both structures convey exhaustivity, clefts and sentences with '*nur* ' ('only', in German) do not do so for the same reasons. Apparently, the exhaustivity effect of '*nur* ' is a logical effect, related to the truth values of the sentence, while the exhaustivity conveyed by clefts would be due to the default interpretation of this type of structure as exhaustive, that is, due to the conversational implicature.

In this ERP experiment, the N400 pattern happened in exhaustiveness violations for clefts sentences and the P600 pattern was observed for the same violations in sentences with the adverb '*nur'* ('only'), in German. P600 is a pattern presented in positive waves, which appears between 600ms and 900ms after a critical segment (Friederici; Hahne; Saddy, 2002). Diversely from the N400, the P600 is associated with syntactic reanalysis and complexity, anomaly or logical contradiction (Drenhaus, Zimmermann, and Vasishth, 2010). Examples of P600 elicitation are experiments in which readers expect to find words from one syntactic category but find words from another (*e.g*., in '**Alice knows how to baked cakes'*, one predicts a verb or a noun and finds an adjective). The P600 is elicited by ungrammatical or just infrequent sentences. These results support the idea that exhaustiveness effects for each of these two types of sentences probably occur for different reasons.

In sum, focus adverbs like 'only' and 'só' might trigger exhaustivity and have measurable effects on processing. Their study has revealed the complex interplay between syntax, prosody, semantics, and context, as well as the role of prediction in focus resolution. These findings provide a rich framework for investigating how linguistic structures guide attention, restrict alternatives, and shape sentence meaning across languages.

3.2. Clefts in English and in BP

According to Braga and De Mello Barbosa (2009 *apud* Alves *et al*., 2015), BP allows different types of clefts (*clivadas*). *Clivadas totais* are the most natural and frequent[8] and they constitute the core of this research. *Clivadas totais* correspond, syntactically and semantically, to the *it-cleft* in English[9], which are included in this project. This project also aims in

[8]"It was MARIA who brought the book": example taken from Alves *et al.* (2015).
[9]"*It was MARIA who brought the book* ": free translation of the example by Alves *et al.* (2015) for the English language.

sentences with the adverbs '*só*', in BP, and 'only' in English, interpreted as a type of focus particle, so they are able to: i) focus on a specific part of the sentence; ii) combine with a specific constituent; iii) have a specific semantic scope (König, 1991).

However, there is evidence that exhaustivity in clefts are processed differently to exhaustivity in focus particles: Drenhaus, Zimmermann and Vasishth (2010) have conducted experiments in German with the adverb '*nur*' and Lignani and Fonseca (2023), with 'só', in PB. Research demonstrates a stronger exhaustive effect when adverbs are present, compared to cleftings. Additionally, Drenhaus, Zimmermann and Vasishth (2010) found different ERPs associated with the two different types of sentences. Clefts indicated a semantic analysis process for interpreting exhaustivity and a syntactic exhaustivity mechanism for the focusing adverb '*nur* '. This interpretative distinction in the processing of the two different focus strategies suggests that the prediction of exhaustivity marked by each of them may happen by different strategies. Assuming that exhaustivity in clefts depends on pragmatic (Horn, 1981; Krifka, 2008; Drenhaus *et al.*, 2010), the computation of the exhaustivity is expected to be more sensitive to context information in clefts than in 'only'-sentences.

Cleft constructions, such as it-clefts in English ('It was John who called') and their counterparts in BP ('*Foi o João que ligou'*), are marked syntactic structures that serve to assign focus to a constituent, while shadowing the rest of the sentence. These constructions typically consist of a cleft clause and a clefted element (the focused constituent), followed by a relative clause that conveys the background information. Clefts are considered prototypical focus-marking devices across languages (Lambrecht, 1994; Kiss, 1998; Destruel *et al*., 2015), and they play a crucial role in encoding and signaling contrastive focus.

Semantically, cleft constructions often give rise to an inference of exhaustivity - that is, they imply that no alternative to the clefted element satisfies the proposition described by the relative clause. For instance, in It was Mary who solved the problem, the sentence suggests that Mary, and no one else, solved the problem. This exhaustivity effect has been shown to be stronger and more obligatory in clefts than in canonical sentences, although its exact nature – whether it is a semantic presupposition, a conventional implicature, or a conversational implicature – has been widely debated (Halvorsen, 1978; Horn, 1981; Velleman *et al*., 2012; Büring; Kriz, 2013).

Recent accounts of clefts further support the view that their exhaustivity is not rigidly encoded, but dependent on the focus structure and on the discourse context. Velleman *et al*.

(2012) suggests that the exhaustivity inference associated with clefts must be sensitive to focus. Crucially, the same appears to hold for their existential presupposition. In contexts such as the ones proposed by Onea (2019), with a question under discussion (QUD) "Who attacked whom?", the use of a cleft "It was Donald who attacked Megyn" did not require that it is already established that Megyn was attacked. Rather, the interpretation is constructed relative to the QUD, that is, the context, allowing the sentence to introduce and resolve both thematic roles simultaneously. This suggests that the presuppositional content of clefts is not fixed at the level of the pivot-relative clause structure (at-issue). Instead, it might be dynamically anchored to the focused elements and the set of relevant alternatives. As a result, both the existential presupposition and the exhaustivity inference are context-dependent and comparatively flexible (Onea, 2019).

Several experimental studies have investigated whether clefts are processed as conveying exhaustive meaning and how this affects real-time comprehension. One line of research suggests that clefts incur processing costs due to their structural complexity, but facilitate interpretation in contexts where contrastive focus is expected. For example, Birch and Clifton (1995) showed that readers take longer to process clefts than canonical structures but are more likely to draw exhaustive inferences. Similarly, Destruel and colleagues (2015; 2018) examined it-clefts across several languages and found consistent patterns of exhaustive interpretation, though with variation in strength and obligatoriness depending on language and context.

In the domain of sentence processing, eye-tracking studies have revealed that clefts elicit longer reading times at or shortly after the clefted element, suggesting that comprehenders allocate more attention to this position to resolve focus-related inferences (Cowles, 2003; Cowles et al., 2007; Destruel; De Veaugh-Geiss, 2018). These results support the view that clefts serve as structural cues for focus, guiding the comprehender's allocation of cognitive resources and expectation about the discourse structure.

Cleft constructions are also highly relevant in relation to discourse coherence and information structure. According to Lambrecht (1994), clefts are used to introduce contrastive or corrective focus, particularly in contexts where an alternative has been evoked or questioned. This function makes effective devices for resolving ambiguity, correcting presuppositions, or highlighting mismatches between speaker and listener beliefs. As such,

clefts not only shape sentence-level interpretation but also contribute to larger-scale discourse organization.

In BP, cleft constructions are syntactically and functionally similar to those in English, but they exhibit some specific distributional and prosodic properties. The canonical form '*Foi X quem Y*' mirrors the English '*It was X who Y'*, with the verb 'foi' acting as the copula and a relative clause headed by 'quem' – or 'que', often. Like it-clefts, BP clefts (*clivadas)* typically mark narrow focus and license an exhaustive interpretation of the clefted constituent (Santos, 2002; Alves *et al*., 2015). Experimental evidence suggests that Brazilian speakers interpret clefts as signaling exclusivity, though the strength of this interpretation can be modulated by intonation, discourse context, and presupposition triggers.

Recent cross-linguistic research has attempted to disentangle the factors that influence exhaustivity in clefts across languages. DeVeaugh-Geiss and colleagues (2020) argue that exhaustivity is not uniformly encoded in clefts but results from the interaction between structure, prosody, and context. In particular, they show that languages vary in how obligatory the exhaustivity effect is, and whether it is encoded semantically or pragmatically. BP appears to occupy a middle ground: clefts tend to implicate exhaustivity, but they also allow for contextual accommodation or cancellation under certain discourse conditions.

From a processing standpoint, clefts may impose additional load on the parser due to their biclausal structure and the need to integrate the focused constituent with an embedded proposition. However, they may also reduce ambiguity by making the focus structure explicit, facilitating integration when the discourse requires contrast or correction. This dual effect mirrors findings in the good-enough processing literature (Ferreira; Bailey; Ferraro, 2002), which shows that readers are sensitive to both structural complexity and functional informativeness when interpreting non-canonical constructions.

Furthermore, cleft constructions engage predictive mechanisms during comprehension. Comprehensioners may anticipate that clefted constituents are likely to be contrastive or exclusive, and this anticipation affects the allocation of processing resources (Cowles, 2003). When expectations meet, clefts facilitate integration; when they are violated, readers incur reanalysis costs. This pattern aligns with predictive models of language comprehension (Kuperberg; Jaeger, 2016; Pickering; Gambi, 2018), which posit that comprehensioners have expectations about upcoming linguistic *input* based on probabilistic cues and discourse constraints.

There is evidence that clefts and '*só*'/ 'only'–sentences unleash linguistic predictions related to semantic exhaustiveness (Paterson, Liversedge; Underwood, 1999; Liversedge, Paterson and Clayes, 2002; Paterson *et al*., 2007; Filik, Paterson; Liversedge, 2009; Drenhaus, Zimmermann and Vasishth, 2010; Carlson, 2014; 2015; Lignani; Fonseca, 2023). Our hypothesis is that predictions may occur in different ways, since the exhaustiveness is different for each structure. The violation of exhaustiveness in sentences with the focusing particle 'only' evokes a P600, which indicates syntactic reanalysis. Exhaustiveness violation effects in clefts sentences elicit an N400. So, exhaustiveness in clefts seems to be due to contextual semantic effects (maximum presupposition) or the default pragmatic strengthening of the exhaustive interpretation (Horn, 1981).

Our hypothesis is that prediction-by-production is responsible for the plausibility of the *inputs* in clefts, taking into account the conversational implicature and being responsible for producing the exhaustive effect in the clefting. The optionality of prediction-by-production could be a reason why exhaustiveness violations in clefts structures are more acceptable. Differently, prediction-by-association would be responsible for prediction in sentences with the adverb '*só*'/'only', as it is an automatic and ubiquitous process of widespread activation, generating greater syntactic complexity and triggering processes of reanalysis and updating of the mental linguistic representation, given the violation of exhaustiveness (Drenhaus, Zimmermann and Vasishth, 2010).

4. METHODOLOGY

The present study aims to test the following hypotheses: 1) in English and in BP, exhaustivity violations in it-clefts/*clivadas totais* are more acceptable than the same violations in sentences with the focusing adverb 'only'/*'só*' (Drenhaus, Zimmerman and Vasishth, 2010); 2) (non-)exhaustivity in it-clefts/*clivadas totais* are more sensitive to a context integration than sentences with 'only'/'*só*'; 3) the violation of exhaustivity in it-clefts/*clivadas totais* unleashes a process of semantic cancellation during comprehension, while the same violation in 'only'-sentences might evidence syntactic reanalysis and/or repair.

The method of this research comprehended empirical tasks. We developed five (5) experiments to test our hypotheses: 3 (three) acceptability questionnaires (two in English and one in BP); and 2 (two) experiments with the eye tracking technique (one in in English and one in BP). Except for the eye-tracking experiment in BP, all of them were already applied and their data has been analysed. The following sections describe all of these experiments.

4.1. Acceptability Judgment Tasks (AJTs)

In this Section, we will present the Likert tests (Likert, 1932) – AJTs – that we developed in English (1 and 2) and in BP (3). For the purposes of this experiment, it is important to highlight that our 5-point Likert responses were treated as ordinal data[10].

The first hypothesis we wanted to test was that 1) exhaustivity violations in it-clefts/*clivadas totais* are more acceptable than violations of the same nature in sentences marked with 'only'/'*só*'. As we have discussed in some of the previous sections, there is evidence suggesting that clefts and focus adverbs convey exhaustivity through different processes. Clefts would convey exhaustivity through conversational implicature, and focus adverbs might associate with the truth-functional values of the sentence (Horn, 1981; Kiss, 1998; Drenhaus; Zimmerman; Vasishth, 2010).

We designed two tasks to measure the acceptability of these structures in English (tasks 1 and 2). The design, the method of application and the method of analysis of such

[10] This is because responses corresponded to ordered categories that are non-equivalent and not guaranteed to be equidistant.. Consequently, we did not consider the responses on our scale to be treated as intervals (Joshi *et al*, 2015). Hence, data was analysed using cumulative link mixed models (*clmm*) with random effects, as we will present in the next subsections.

experiments are the same, but task 1 and task 2 differ in their materials: the experimental items in the task 1 followed the pattern: ‘and’ + (subject/object) + ‘too’; for items in the task 2, we used the pattern ‘and also’ + (subject/object). Despite being both additive particles (König, 1991), ‘too’ and ‘also’ differ syntactically, semantically and pragmatically (Grubic *et al*., 2019); consequently, they could influence participants’ perception of (non-)exhaustivity in different ways, which might affect the rates of acceptability of a sentence.

Previous studies show that ‘too’ triggers stronger presupposition demands and signals discourse similarity (Winterstein, 2011). On the other hand, ‘also’ is more neutral and it often comes before the nominal phrase (NP) in written contexts, which requires that the reader recovers a piece of information that was given before (Hüttner; Von Heydebrandt; Wessels, 2004; Höhle *et al*., 2009; Drenhaus; Zimmerman; Vasishth, 2010). Because of that, ‘too’ is more effective to create a syntactic/semantic parallelism than ‘also’, which makes the latter likely to be more difficult to process than the former. Besides, it is important to mention that these particles occupy different positions in a sentence, and this is evidence enough to consider that they are processed differently (Cruschina, 2019). Considering such peculiarities of each additive particle, we tested both kinds of structures and compared one another. As we previously mentioned, we have also developed a third task (task 3) in BP, which differs from the previous two only in relation to the language of the stimuli.

It is worth noting that all acceptability tasks presented stimuli *out-of-the-blue*, that is, isolated sentences with no preceding context. This methodological choice is common in grammaticality and focus-related studies, as it allows researchers to assess participants’ intuitions without any additional discourse cues (Paterson *et al*., 2007; Filik; Paterson; Liversedge, 2009; Carlson, 2014; Drenhaus *et al.*, 2010; Lignani; Fonseca, 2023).

The next subsections discuss the details of the acceptability tasks 1 and 2.

4.1.1. Acceptability Task 1 – AJT1 (in English)

The first experiment that we will describe in this paper is the first Likert test that we applied in English, with the additive particle ‘too’.

4.1.1.1. Materials

We developed 48 sets of experimental sentences[11]. All of the experimental sentences had the following structure: temporal adverbial + (‘only’/it-cleft) subject-NP + transitive verb + ‘and…did too’.

There were 48 fillers (24 grammatical and 24 ungrammatical[12]) and 3 sentences of practice displayed one by one, before the actual trial. Among the experimental items, half of the subject-NPs (24) were common nouns[13] and the other half were proper nouns (Huddleston; Pullum, 2002). Within the common nouns, there were 14 gender-neutral, 5 feminine and 5 masculine, whereas among the 24 subject-NPs proper nouns, 12 were feminine and 12 were masculine. We made this with the purpose to minimise gender biases.

Each set contained a sentence in 4 (four) different conditions: a) with the adverb 'only', conveying exhaustivity ([+exh]); b) with the adverb 'only’, but non-exhaustive ([-exh]); c) an it-cleft, conveying exhaustivity ([+exh]); and d) an it-cleft, but non-exhaustive ([-exh]).

For example, the set number (7) below:

(7) a. At Christmas, **only the grandmother** prepared the meals and the drinks too. ([+exh])

b. At Christmas, **only the grandmother** prepared the meals and **the grandfather** did too. ([-exh])

c. At Christmas, **it was the grandmother who** prepared the meals and the drinks too. ([+exh])

d. At Christmas, **it was the grandmother who** prepared the meals and **the grandfather** did too. ([-exh])

[11] See Appendix A for all the items.

[12] We used ungrammatical items, such as ‘**In the morning, she and I gone to the park’*, in this experiment with two purposes. The first one was to guarantee that participants were unaware about what we were testing. The non-exhaustive conditions represented a violation and, subsequently, lower rates of acceptability. So, participants could easily become suspicious about the experimental items, if those were the only unacceptable/weird sentences in the whole trial. The second reason for this choice of ours is described better in our 4.1.1.6 Results and Data Analysis subsection. In sum, we used ungrammatical fillers as a parameter of participants’ attention during the task, since all of them took the task remotely. We could not watch participants doing the task, so we wanted to take extra precautions about some amount of their distraction – and results’ reliability – levels.

[13] These were all animated common nouns for people, divided between family/relationships (*e.g.*: grandmother, father) and occupations (*e.g.*: doctor, teacher).

In this example, the object-NP '*the drinks*' did not offend the exhaustivity expected by '*only*', because it did not conflict with the idea that the grandmother is the one person who prepares both meals and drinks. Therefore, sentences (7a) and (7c) are exhaustive.

On the other hand, '*the grandfather*', a subject-NP for the same action ('preparing the meals'), represents a disruption to the exclusiveness regarding '*the grandmother*' being the only person who prepares meals at Christmas. Because of that, sentences (7b) and (7d) are non-exhaustive, as they present a violation of the exhaustivity that people would expect by reading 'only'-sentences or it-clefts.

Importantly, the stimuli comprehended these two types of structure – 'only'-sentences (7a; 7b) and it-clefts (7c; 7d)). As we wanted to observe the effect of (non-) exhaustivity in each one of them, we had exhaustive (7a; 7c) and non-exhaustive (7b; 7d) sentences for each syntactic structure.

4.1.1.2. Method

This experiment was designed on the website PC IBEX (Zehr; Schwarz, 2018). The website provided us with a sharable link for data collection and we shared the link on the SONA SYSTEMS[14] (2024) interface, so that people could participate on the trial remotely. We built the trial using the Latin square *design,* with four versions of the experiment (A, B, C and D), which means that each participant could see only one condition from each experimental set, and the same number of items per condition [15](Abbuhl; Gass; Mackey, 2013). PC IBEX (Zehr; Schwarz, 2018) randomised items and versions of the experiment automatically, every time a new participant joined the trial. Once participants accessed the link provided, they saw an initial page (Figure 1):

---

[14] This is an online system used for participant recruitment and study management. This System is used in the School of Psychology at the University of Leicester (UoL), where we gathered participants for all the experiments in English. As we will discuss in section *4.1.1.4 Participants*, people who participated in this task were all students of Psychology at the UoL, so they signed up for experiments on SONA, and received credits for their participation.

[15] For example, participant P1 sees version A of the experiment. At the end of the experiment, the Latin square design guarantees that P1 will have read the same number of sentences for all levels of condition (a), (b), (c), (d), and only one condition per experimental set. In our case (conditions = 4; items = 48), P1 reads each condition 12 times, one from each experimental set (*e.g.*: the exhaustive condition with 'only' is read in: 1st, 5th, 9th, 13th, 17th, 21st, 25th, 29th,33rd, 37th,41st, 45th; and this pattern repeats across all the other 3 levels of conditions).

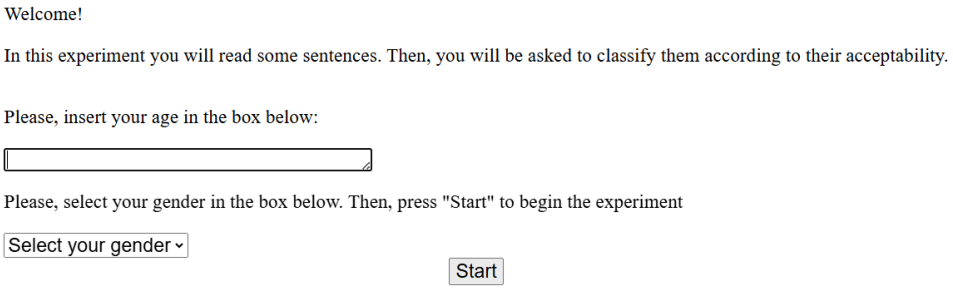


Figure 1 – Welcome screen of the acceptability tasks in English.

After filling the forms with the information requested, participants read the instructions on the next page (Figure 2):

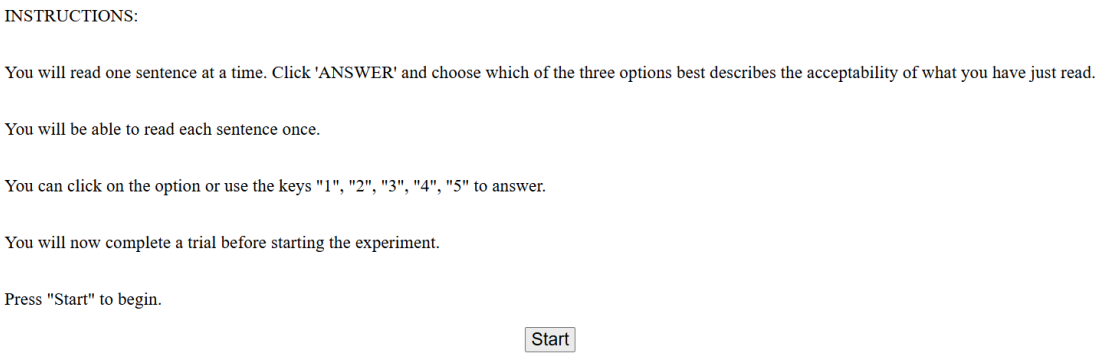


Figure 2 – Instructions of the acceptability tasks in English.

Then, participants pressed a button to start a trial with 3 practice-sentences (Figure 3) before they read the experimental items. Each sentence was shown on the screen, on top of an '*Answer*' button, which participants should select when they were ready to answer the question and read the next sentence.

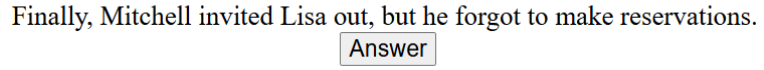


Figure 3 – Example of practice-sentence of the acceptability tasks in English.

After having read a sentence and clicked the button '*Answer*' (Figure3), participants should answer the following question (Figure 4):

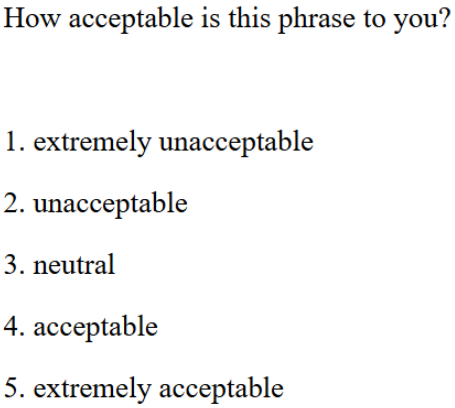


Figure 4 – Acceptability scale of the acceptability tasks in English.

They could answer the question by pressing the correspondent keys on their keyboards (1, 2, 3, 4 or 5) or by selecting their answer with their mouse. Once they completed the 3-sentence practice, they were ready to begin the experiment (Figure 5):

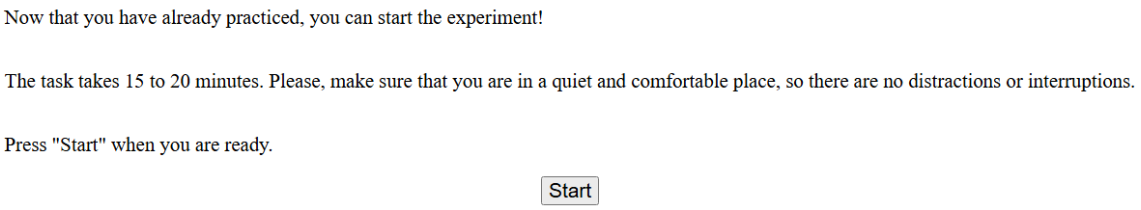


Figure 5 – Beginning of the experiment of the acceptability tasks in English.

The experimental items, such as the example in Figure 6, were randomised and displayed one by one on the screen. After each sentence, participants should answer the acceptability question in Figure 4, just as they did during the practice phase:

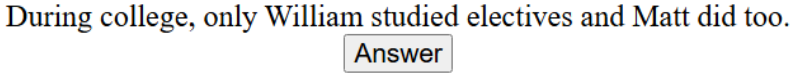


Figure 6 – Example of an experimental item of the acceptability task 1 in English.

After the participants have gone through the 99 sentences (48 experimental ones, 48 fillers, plus 3 practice-sentences), a message appeared on the screen to let them know that the

trial was over. On the last screen, there was a link to the SONA SYSTEMS (2024), so their participation was computed and they received credits for it.

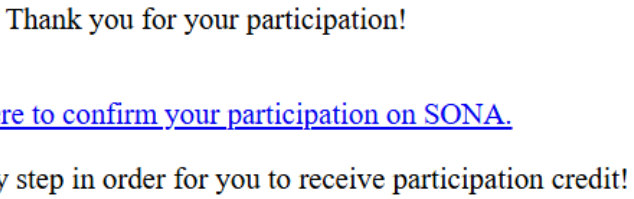


Figure 7 – Final screen of the acceptability tasks in English.

4.1.1.3. Independent Variables

As we manipulated the semantic constituent of exhaustivity, the latter was the independent variable (IV) of this experiment. This happened by choosing the nature of the critical segment after the conjunction '*and*' in the clause, which could be either a new object-NP or a new subject-NP for the same verb. That essentially means that this constituent could convey exhaustivity or not. We had two types of structure ('only'-sentences and it-clefts) to test exhaustivity, and so the second IV of this task was the syntactic structure..

4.1.1.4. Dependent Variables

The question that the participants should answer was: "How acceptable is this phrase for you?". Participants were supposed to rate the sentences, according to the scale: 1. extremely unacceptable; 2. unacceptable; 3. neutral; 4. acceptable; or 5. extremely acceptable. The dependent variables (DVs) of this experiment were the different marks on a 1-to-5-scale of acceptability (Figure 4), and the reading times (TR_reading) of the sentences, which we used primarily as a parameter to identify outliers.

4.1.1.5. Participants

Fifty-six (56) undergraduate students in the School of Psychology of the University of Leicester (England) participated in this experiment. Before initiating the data analysis, we balanced the table of results and remained with data from 44 people, 11 participants for each version of the experiment (A, B, C or D). All of the participants were native English speakers from the United Kingdom (UK), with ages between 18 and 20 years old. The mean age was 19.07 years, and 38 of them were women, 2 men and 2 of them did not select any gender.

Participants (n=44; M= 19.07; 38 women) took the task in exchange for obligatory credits in participation of experiments, which is a graduating requirement of the School of Psychology. They applied to take the experiment on the SONA SYSTEM (2024) and we inserted a link for the initial page of this task (1) on SONA SYSTEM.

4.1.1.6. Results and data analysis - task 1

The data of this experiment was analysed on *RStudio* (R Core Team, 2025). We started our data analysis by using two parameters to remove outliers: the accuracy rate of (un)acceptability for ungrammatical fillers, and the amount of time that participants took to read the experimental sentences.

It is worth recalling that we added ungrammatical 24 fillers (50% of the fillers), such as '*Max and Danny him don't like to pizza*' (filler 82Da). Participants who were attentive during the task were expected to rate ungrammatical fillers as unacceptable sentences. Considering a 1 to 5 scale of acceptability in which 5 is the most acceptable, this means that these sentences would be rated as 1 (extremely unacceptable) or 2 (unacceptable).

As this was a task executed remotely, we believe that this extra attention-control tool was useful to address possible issues about the reliability of the results, once ungrammatical fillers rated as neutral (3) or acceptable (4 or 5) probably meant distractions or disruptions[16]. Hence, we discarded the data of the participants who were not able to rate at least 80% of the ungrammatical items as unacceptable (1 or 2) to balance the versions of the experiment. By the time we proceeded to the data analysis, we only had data from people who were evidenced to be attentive throughout the trial.

After balancing the results table, we generated boxplots and a histogram for the time that participants took to read the sentences, namely TR_reading. We wanted to make sure that we removed outliers before starting our inferential analysis. TR_reading was the measure of the time between the moment that a sentence was shown on the screen and the moment that participants pressed the '*Answer*' button to rate that sentence. The boxplot and the histogram

[16] Although we thought this was probably a good idea for concentration-control issues, later discussions brought our attention to how the insertion of ungrammatical items might have been a controversial choice considering the experimental design itself. That is because if we take into account that participants' task was to classify sentences judging acceptability/adequacy, the presence of declared ungrammatical sentences would potentially lower the mean of acceptability/adequacy, as those were supposedly judged as the least acceptable/adequate.

showed that we had some outliers, and that the data did not follow a normal distribution, as we can see in Figures 8 and 9:

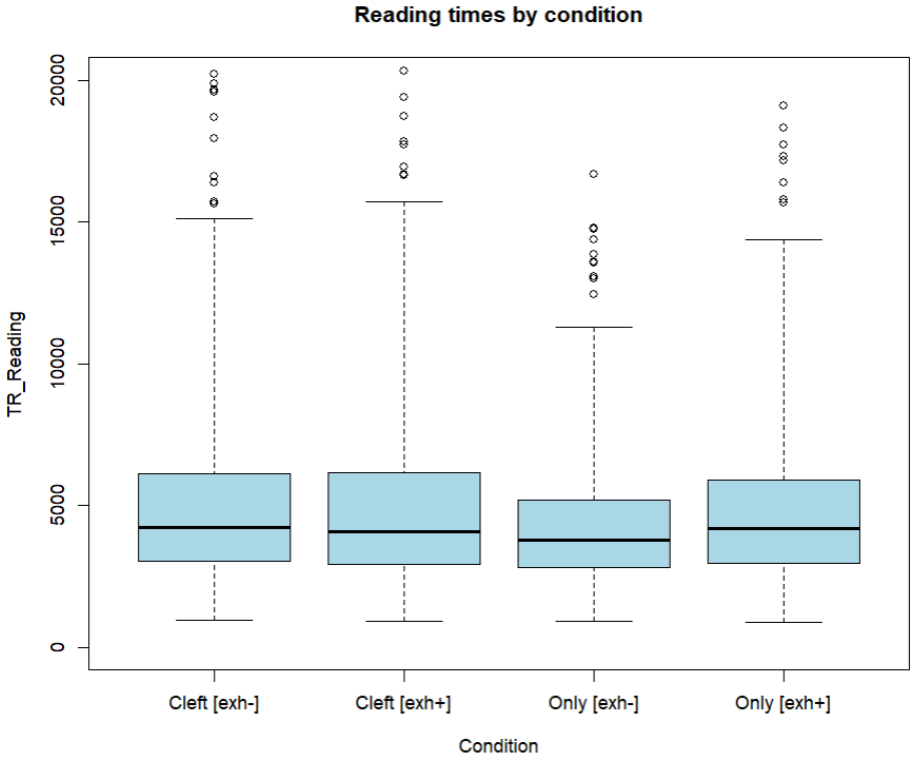


Figure 8 – Boxplot of TR_reading per condition in AJT1.

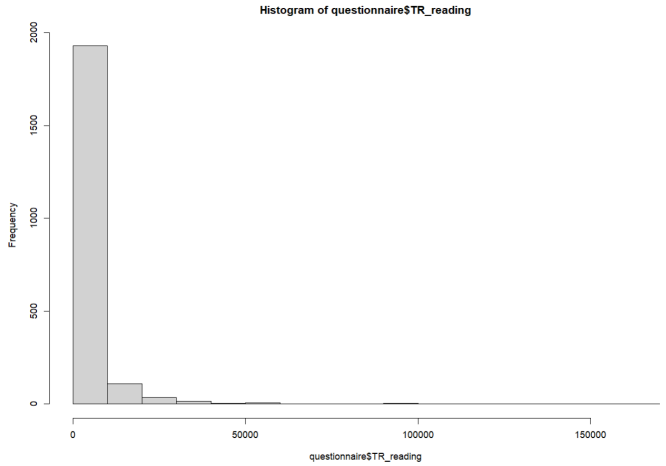


Figure 9 – Histogram of AJT1 before normalisation.

After having a descriptive analysis of the TR_reading measures, we applied a filter of 1.000ms < TR_reading < 15.700ms. Then, we deleted the outliers and we ran a *Lilliefors* (*Kolmogorov-Smirnof*) normality test, which indicated that the distribution of log-transformed

reading times did not significantly deviate from normality (D = 0.020, p= 0.05388). The histogram below shows data after normalization (Figure 10):

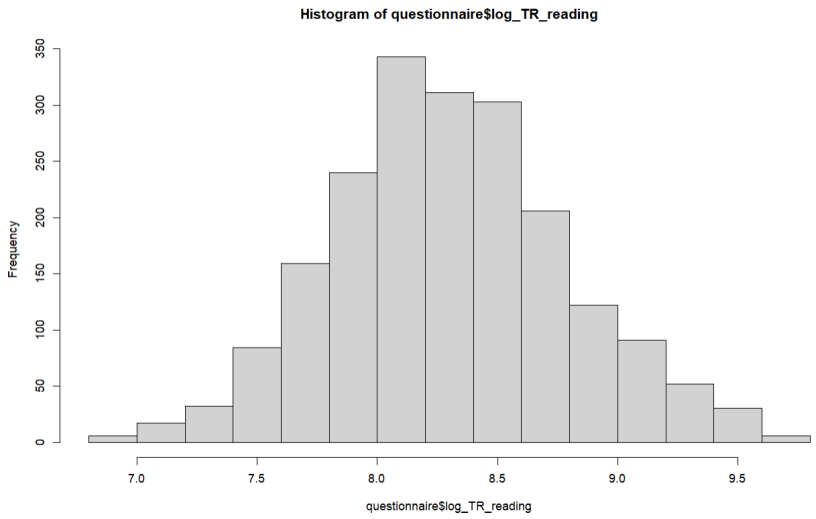


Figure 10 – Histogram of AJT1 after normalisation.

The former results table contained 2112 observations, and after the removal of the outliers, we were left with 2002 observations. We modeled a linear mixed-effects regression (LMM) using the function *lmer* on *RStudio* to check if there was a difference between the reading times (TR_reading) across conditions. This would evidence whether any conditions were less/more rapidly processed than others.

Non-exhaustive 'only'-sentences had the shortest mean of reading time, in comparison to the other conditions, as shown in (Figure 11):

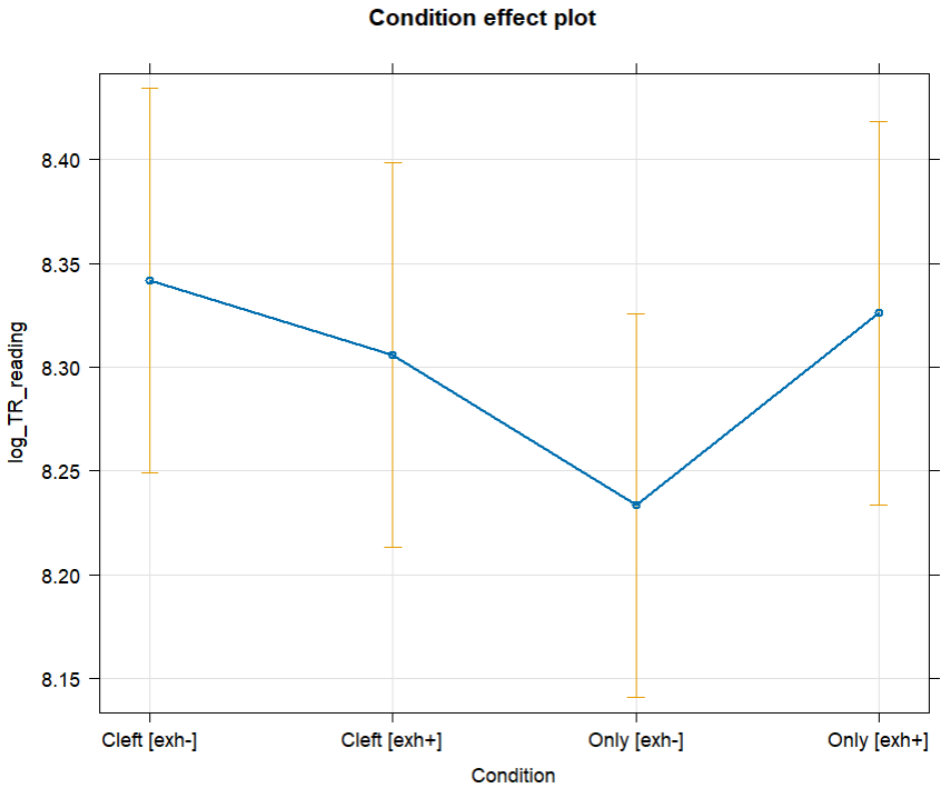


Figure 11 - Log reading times across conditions in AJT1.

The first LMM that we analysed (Table 1) revealed a significant effect of exhaustivity and structure, when these variables were combined: the focus particle 'only' and non-exhaustivity reduced the reading times when they were combined. In other words, the effect of exhaustivity depended on the structure. Variance was estimated for participants (τ_{00} = 0.08) and items (τ_{00} = 0.01):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.31	8.21 – 8.40	**<0.001**
Structure [Only-sentence]	0.02	-0.04 – 0.08	0.487
Exhaustivity [Non_exhaustive]	0.04	-0.02 – 0.09	0.218
Structure [Only-sentence] × Exhaustivity [Non_exhaustive]	-0.13	-0.21 – -0.05	**0.002**
Random Effects			
σ^2	0.16		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.08		
ICC	0.35		
$N_{\ Participant}$	44		
$N_{\ Item}$	192		
Observations	2002		
Marginal R^2 / Conditional R^2	0.007 / 0.353		

Table 1 – LMM for reading times in AJT1.

A new model (Table 2), which compared conditions between one another, made it easier to notice such effect of reading times in non-exhaustive 'only'-sentences (95% CI: -0.17 to -0.05, $p < .001$):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.34	8.25 – 8.43	**<0.001**
Condition [Cleft [exh+]]	-0.04	-0.09 – 0.02	0.218
Condition [Only [exh-]]	-0.11	-0.17 – -0.05	**<0.001**
Condition [Only [exh+]]	-0.02	-0.07 – 0.04	0.588
Random Effects			
σ^2	0.16		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.08		
ICC	0.35		
$N_{\ Participant}$	44		
$N_{\ Item}$	192		
Observations	2002		
Marginal R^2 / Conditional R^2	0.007 / 0.353		

Table 2 – LMM for reading times comparing conditions in AJT1.

It is important to note the differences in the length of the presented stimuli make our findings even more unexpected. Let us take for example the set of conditions below (8):

(8) a. Yesterday, **only Alice** made coffee and tea too. ([+exh])
b. Yesterday, **only Alice** made coffee and **Lidia** did too. ([-exh])
c. Yesterday, **it was Alice who** made coffee and tea too. ([+exh])
d. Yesterday, **it was Alice who** made coffee and **Lidia** did too. ([-exh])

Comparing it-clefts with 'only'-sentences, the first ones were longer than the later, and so exhaustive clefts (8c) were longer than exhaustive 'only'-sentences (8a) and non-exhaustive clefts (8d) were longer than non-exhaustive 'only'-sentences (8b). That was because it-clefts' syntactic structure demands more lexical items ('*it was*/*is*'; '*who*'), in comparison to focus particle ('only') sentences.

Non-exhaustive clefts (8d) were longer than the exhaustive ones (8c), since the former included the auxiliary verb '*did*'/'*did*'. However, we learned that there was no significant difference between reading times for exhaustive or non-exhaustive clefts (Table 4):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.34	8.25 – 8.43	**<0.001**
Condition [Cleft [exh+]]	-0.04	-0.09 – 0.02	0.216
Random Effects			
σ^2	0.17		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participant}$	0.08		
ICC	0.33		
$N_{\ Participant}$	44		
$N_{\ Item}$	96		
Observations	992		
Marginal R^2 / Conditional R^2	0.001 / 0.328		

Table 3 – LMM for reading times for it-clefts in AJT1.

Differently to what happened with clefts, non-exhaustivity for 'only'-sentences influenced participants' reading times (Table 4), and this happened even though non-exhaustive 'only'-sentences were longer than their exhaustive counterparts:

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.23	8.14 – 8.33	**<0.001**
Condition [Only [exh+]]	0.09	0.04 – 0.15	**0.001**
Random Effects			
σ^2	0.15		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.08		
ICC	0.36		
$N_{\ Participant}$	44		
$N_{\ Item}$	96		
Observations	1010		
Marginal R^2 / Conditional R^2	0.009 / 0.362		

Table 4 – LMM for reading times for 'only'-sentences in ATJ1.

Considering that the stimuli in the non-exhaustive conditions corresponded to the longest syntactic structures, it is interesting to notice that non-exhaustive 'only'-sentences (Only [exh-]) had the shortest reading times. This suggests that reading times for 'only'-sentences was not necessarily associated with the length of the clauses.

At this point, it is interesting to point out that we called 'reading times' (TR_reading) the elapsed time between stimulus onset (presentation on screen) and recorded response (pressing the '*Answer*' button). In other words, the TR_reading can be interpreted as the time that participants' would take to read a sentence and feel ready to make an acceptability judgment for it[17].

Recalling our hypothesis – that is, non-exhaustive 'only'-sentences are less acceptable than non-exhaustive it-clefts – we expected longer reading time means for non-exhaustive 'only'-sentences, as a sign of more complex processing. Nevertheless, the results suggest that non-exhaustive 'only'-sentences were processed differently from other conditions.

[17] We consider that participants' decision about acceptability for each sentence is comprehended in TR_reading time. That is because we argue that participants would press the '*Answer'* button only when they felt ready to rate the sentence, which implies making a decision about what key (1, 2, 3 4 or 5) to press.

After having observed what happened to the reading time across conditions, we proceeded to a descriptive analysis of the acceptability rates. Figure 12 illustrates our findings:

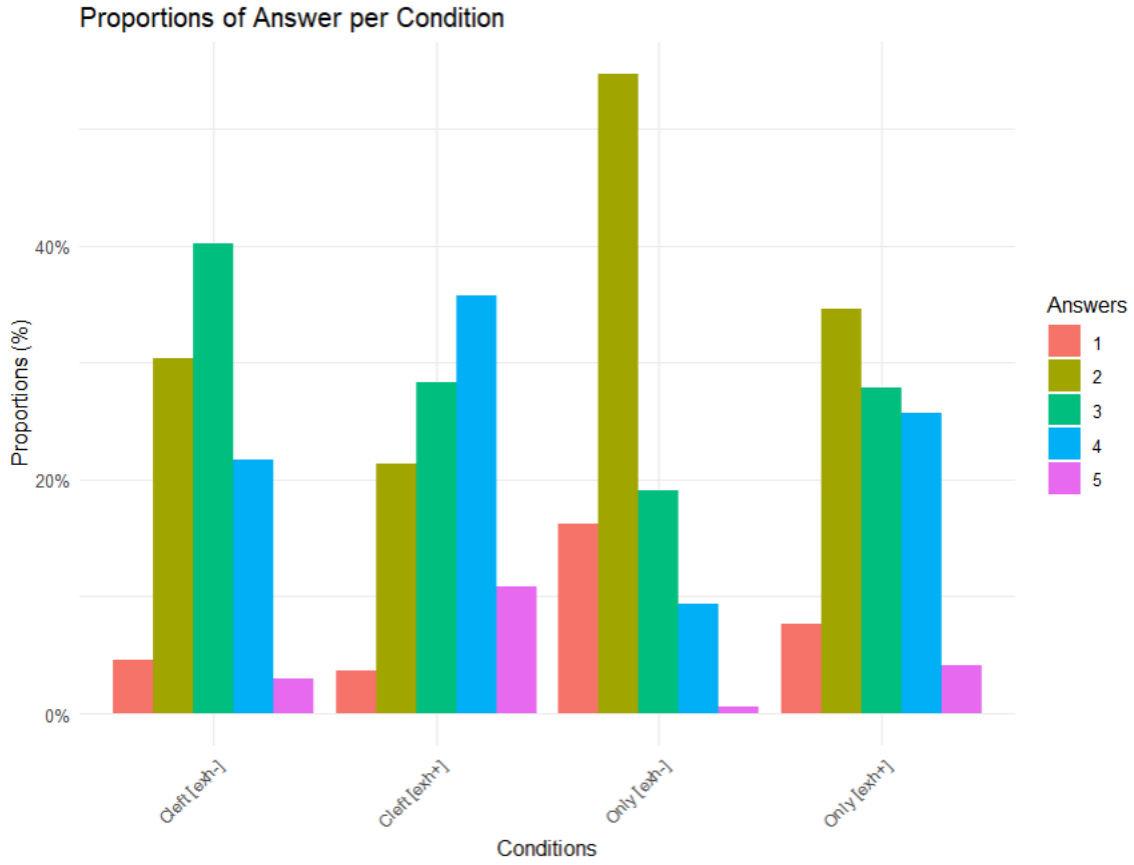


Figure 12 – Acceptability rating across conditions in AJT1.

As we expected, the non-exhaustive condition for 'only'-sentences (Only [exh-]) received the highest percentages of answers 1 and 2 (approximately 70%) of participants' choices, versus 19% neutral (3) and only 10% acceptable (4 and 5).

We have also predicted that the exhaustive condition of the clefts (Cleft [exh+]) would be considered the most acceptable one, for two reasons: firstly, because it did not involve a violation of exhaustivity; secondly, because it would convey exhaustivity due to pragmatic implications, which would probably make it more compatible to additive particles, in comparison to 'only'-sentences. We have seen that, indeed, 46.8% of participants rated exhaustive clefts to be acceptable (answers 4 and 5), versus 28.2% neutral (3) and 25%

unacceptable (1 and 2) ratings. Interestingly, an unexpected pattern emerged when comparing non-exhaustive clefts and exhaustive ‘only’-sentences.

Participants showed a relative preference for non-exhaustive clefts (35% unacceptable; 40% neutral; 25% acceptable) over exhaustive ‘only’-sentences (42.3% unacceptable; 27.9% neutral; 29.8% acceptable). One possible interpretation is that the presence of ‘only’ encodes exclusivity and so inhibits additive continuations (*e.g.*, “…did/did too”), leading to lower acceptability regardless of exhaustivity condition. In contrast, it-clefts, which do not encode exhaustivity syntactically, may allow for more flexible interpretations, even in contexts that could be construed as non-exhaustive.

Notably, the non-exhaustive cleft condition elicited the highest proportion of neutral responses (40%), which may reflect increased uncertainty rather than categorical acceptance or rejection. This pattern suggests that participants did not consistently converge on a single interpretation for these stimuli, in contrast to the more decisive responses observed for ‘only’-sentences.

To provide an initial overview of response distributions, we conducted Pearson’s chi-squared tests: Likert-scale responses were grouped into three categories (unacceptable: 1–2; neutral: 3; acceptable: 4–5) and we tested associations between each variable with the responses. The results indicated significant associations both for syntactic structure ($\chi^2(2) = 150.77$, $p < 0.001$ – Figure 13) and exhaustivity ($\chi^2(2) = 96.26$, $p < 0.001$ – Figure 14). Nevertheless, chi-squared tests do not account for the random effects (variability across participants and items), so these analyses were treated merely as exploratory.

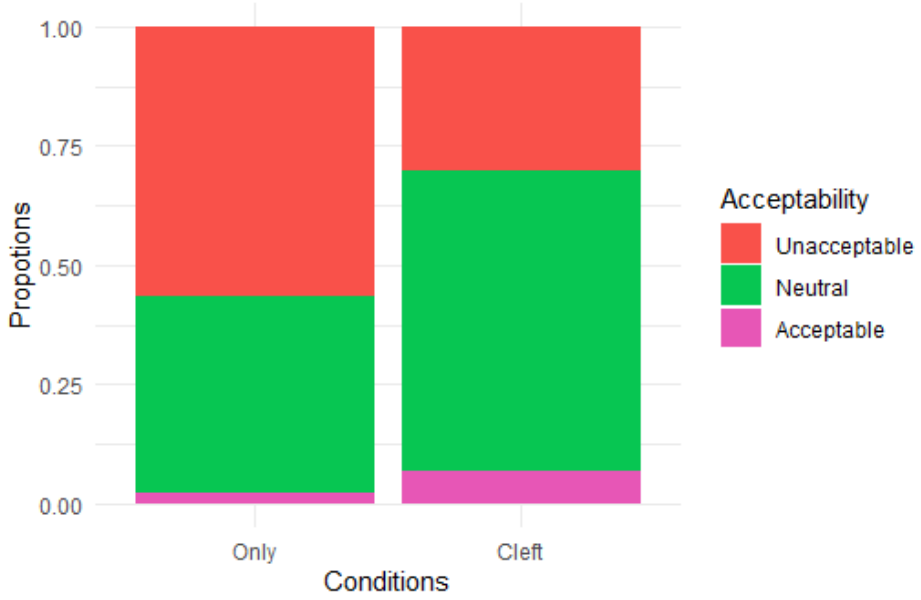


Figure 13 – Proportions of acceptability ratings across syntactic structure in AJT 1.

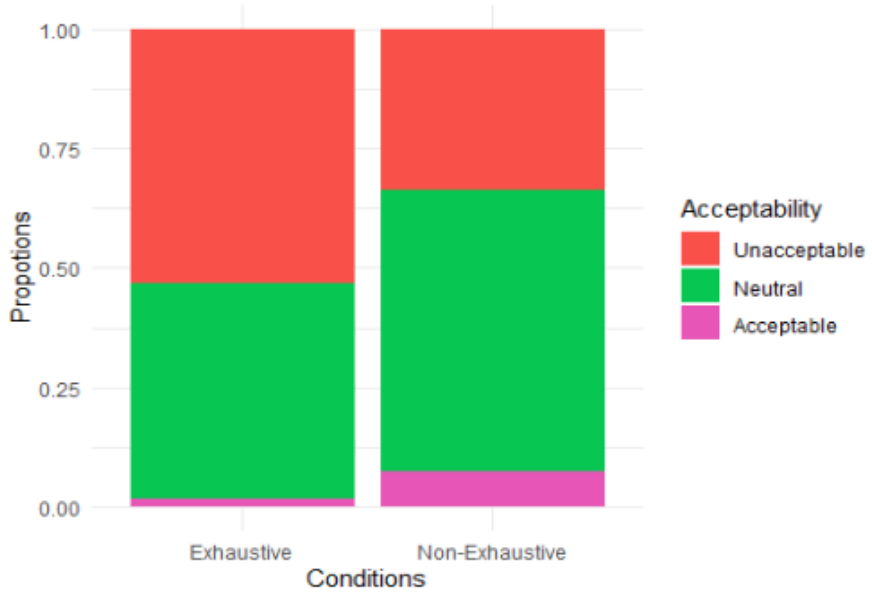


Figure 14 – Proportions of acceptability ratings across exhaustivity in AJT1.

The most important test that we developed was a cumulative-link mixed model (CLMM; function *clmm* on *RStudio*) to examine the effects of syntactic structure and exhaustivity on acceptability ratings. This model estimates the cumulative log-odds of a higher response on the acceptability scale (DVs), in relation to the syntactic structure and the exhaustivity (IVs). The DV variable was treated as an ordered factor ranging from 1 (extremely unacceptable) to 5 (extremely acceptable), and the model included the interaction

of both IVs (syntactic structure and exhaustivity), as fixed effects. We included intercepts for participants and items as random effects.

The results presented odds ratios with 95% confidence intervals (Table 5). 'Only'-sentences (OR = 0.33, 95% CI [0.23, 0.47], $p < 0.001$) had a negative effect in acceptability, and so did non-exhaustivity (OR = 0.38, 95% CI [0.27- 0.54], $p < 0.001$). Notably, there was a minor negative interaction (OR = 0.54, 95% CI [0.33, 0.89], $p = 0.016$) between structure and non-exhaustivity. Thresholds for cumulative probabilities between response categories also showed expected progression (e.g., for 3|4 OR = 1.33, 95% CI [0.82, 2.17], $p = 0.246$; 4|5 = 19.13, 95% CI [11.36, 32.22], $p < 0.001$). The model included random effect variances for items (τ_{00} = 0.42) and participants (τ_{00} = 1.95), with an interclass correlation coefficient (ICC) of 0.42.

These findings suggest that the effect of exhaustivity violations on acceptability judgments is not uniform across structures, which matches our previous findings. While exhaustive clefts tended to receive relatively high ratings, non-exhaustive 'only'-sentences were generally judged less acceptable, as this condition ('Only' [exh-]) presented the interaction of variables that decreased acceptability.

	Answer		
Predictors	*Odds Ratios*	*CI*	*p*
1|2	0.01	0.00 – 0.01	**<0.001**
2|3	0.21	0.13 – 0.34	**<0.001**
3|4	1.33	0.82 – 2.17	0.246
4|5	19.13	11.36 – 32.22	**<0.001**
Structure [Only-sentence]	0.33	0.23 – 0.47	**<0.001**
Exhaustivity [Non_exhaustive]	0.38	0.27 – 0.54	**<0.001**
Structure [Only-sentence] × Exhaustivity [Non_exhaustive]	0.54	0.33 – 0.89	**0.016**
Random Effects			
σ^2	3.29		
$\tau_{00\ Item}$	0.42		
$\tau_{00\ Participant}$	1.95		
ICC	0.42		
$N_{Participant}$	44		
N_{Item}	192		
Observations	2002		
Marginal R^2 / Conditional R^2	0.142 / 0.501		

Table 5 – CLMM with interaction of the variables in AJT1.

As we can see in Table 5, the non-exhaustive conditions had a negative effect, which means that the non-exhaustivity presumably decreased the odds for participants to choose higher rates of acceptability. That effect was expected, as we predicted that the violations of exhaustivity would probably cause disruptions.

In conclusion, what we observed for AJT1 was that: 1) ‘only’ decreased the likelihood of acceptability; 2) non-exhaustivity tended to decrease the acceptability; 3) ‘only’ and non-exhaustivity combined also contributed for the sentences to be judged less acceptable, but in a minor scale than when such variables were isolated.

4.1.2. Acceptability Task 2 – AJT2 (in English)

We developed the AJT2 after AJT1. As mentioned in section 4.1, the design and method for both acceptability tasks were the same.

There were the same 48 fillers (24 grammatical and 24 ungrammatical) and 3 practicing sentences from task 1, as well as the proportion of 50/50 for female/male subject-NPs, for the 24 proper nouns, and 14 gender-neutral, 5 feminine and 5 masculine.

4.1.2.1. Materials

The differences between tasks 1 and 2 concerned the additive structure of the experimental items: instead of 'and… too', the items in task 2 presented the following structure: temporal adverbial + ('only'/it-cleft) subject-NP + transitive verb + 'and also… (subject-NP/object-NP constituent)'.

Each set contained a sentence that varied across 4 (four) different conditions: a) adverb 'only', conveying exhaustivity ([+exh]); b) adverb 'only', but non-exhaustive ([-exh]); c) an it-cleft, conveying exhaustivity ([+exh]); and d) an it-cleft, but non-exhaustive ([-exh]).

For example, the set number (9) below[18]:

(9) a. At Christmas, **only** the grandmother prepared the meals and also the drinks. ([+exh])

b. At Christmas, **only** the grandmother prepared the meals and also the grandfather. ([-exh])

c. At Christmas, **it was** the grandmother **who** prepared the meals and also the drinks. ([+exh])

d. At Christmas, **it was** the grandmother **who** prepared the meals and also the grandfather. ([-exh])

As in the acceptability task 1, sentences (9a) and (9c) are so-called exhaustive because they respect the expected exclusiveness to the subject-NP. For the opposite reason, sentences (9b) and (9c) are non-exhaustive. For non-exhaustive conditions, it is important to consider what the absence of the auxiliary verb 'do' triggers in this case.

[18] See Appendix B for all the items.

During the acceptability task 1, participants read sentences such as (7b) ‘At Christmas, only the grandmother prepares the meals and the grandfather did too’. In this kind of sentence, there is a VP-ellipsis (‘did too’), which omits ‘the meals’ (object-NP) from the sentence (Roberts; Matsuo; Duffield, 2013). However, in (9b) ‘At Christmas, only the grandmother prepares the meals and also the grandfather’ not only ‘*the meals*’ (object-NP), but also the verb itself (‘*prepares*’) are omitted. These omissions leave us with a structure of coordination and *gapping* (Frazier; Clifton, 2001; Johnson, 2009), that is different from an ellipsis.

Cases of *gapping* processing demand strict coordination and a structural parallelism that depends on additional cost for the *parser*: comprehenders must be able to access previous information to build the meaning of a coordinate sentence with a gap (Johnson, 2009). According to Johnson (2009), additive constructions like ‘and also…’ would be instances of gapping derived via across-the-board (ATB) movement. In ATB, there is no post-syntactically elision of the verb as in an ellipsis, but a simultaneous extraction of the verb from both sentences. This mechanism predicts stricter constraints on parallelism, and may incur higher integration costs during online processing, compared to auxiliary-driven VP-ellipsis. In other words, processing a *gapping* would be more complex than processing a VP-ellipsis.

Furthermore, the absence of a verb makes sentences more susceptible to ambiguous processing. For example, in (9b) ‘At Christmas, only the grandmother prepares the meals and also the grandfather’, the interpretation ‘the grandmother prepares the meals and prepares ‘the grandfather’, although semantically unusual - and unexpected - it is syntactically possible. However, in (7b) ‘the grandfather did too’ there is no flexibility for other interpretations.

Nonetheless, despite their semantic and pragmatic similarities, it becomes clear that ‘too’ and ‘also’ are very different additive particles in terms of syntax and processing. We have demonstrated that there is evidence to believe that ‘and also…’ structures are more difficult to process/understand than ‘… did/did too’ ones. That is because ‘and also…’ depends on *gapping*, which elicits a process that is essentially more complex than VP-ellipsis. In addition ‘and also…’ structures allow ambiguous interpretation, whereas this cannot happen with sentences that end in ‘…did too’.

Due to all of this, we expected to see results that match the higher level of complexity of the items in acceptability task 2, compared to task 1. Our hypothesis still was that 1) in English, exhaustivity violations in it-clefts are more acceptable than exhaustivity violations in

sentences with the adverb ‘only’. Even so, due to the complex nature of ‘and also…’ sentences, we expected a milder effect of exhaustivity than the one we found in task 1.

4.1.2.2. Method

This experiment (acceptability task 2) replicated all the characteristics of the acceptability task 1: it was designed on PC IBEX (Zehr; Schwarz, 2018) and the link for this task became available for participants on the SONA SYSTEMS (2024) interface, so that people could participate on the trial remotely. We kept the Latin square *design,* with four versions of the experiment (A, B, C and D). PC IBEX randomised items and versions of the trial automatically. Once participants accessed the link provided, they saw the pages on Figures 1, 2, 3, 4, 5 and 7[19].

Figure 15 shows an example of sentence displayed during the AJT2:

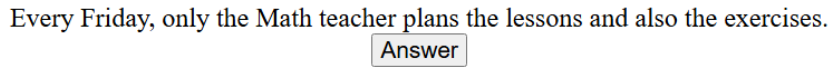


Figure 15 – Example of an experimental item of the AJT2.

The last screen of the experiment, as in task 1, displayed a link to the SONA SYSTEMS (2024), so students could guarantee their credits for participation.

4.1.2.3. Independent Variables

The independent variables, as in task 1, were exhaustivity and the syntactic structure of the items. It is worth noting that the items of task 2 had the ‘and also…’ structure, while the items in task 1 had ‘and… too’. Despite that, items varied equally across conditions and the independent variables were the same for both experiments.

[19] Figure 6 contains an example of sentence displayed during the acceptability task 1.

4.1.2.4. Dependent Variables

The dependent variables (DVs) of this experiment were, once again, the different marks on a 1-to-5-scale of acceptability (Figure 4) and the reading time of the stimuli (TR_reading).

4.1.2.5. Participants

Seventy-two (72) undergraduate students in the School of Psychology of the University of Leicester (UK) participated in this experiment. Before initiating the data analysis, we balanced the table of results and remained with data from 44 people, 11 participants for each version of the experiment (A, B, C or D). All of the participants were native English speakers from the United Kingdom. The mean age was 19 years, and 38 of them were women, 5 men and 1 of them did not select any gender.

Participants (n= 44; M= 19; 38 women) took the task in exchange for obligatory credits in participation of experiments, which is a requirement of the School of Psychology. They applied to take the experiment on the SONA SYSTEM (2024) and we provided this System with a link that led to the initial page of the acceptability task 2. Importantly, we made sure that people who had participated in the previous task (1) were unable to sign up for task 2, so there would be no influence from task 1 in the results of task 2.

4.1.2.6. Results and Data Analysis

The data of this experiment, as of the previous one, was analysed on *RStudio* (R Core Team, 2025). The method of analysis was also very similar to the method of analysis of the previous acceptability task. Again, before manipulating the data, we removed from the results' table participants that did not rate at least 80% of the ungrammatical fillers as unacceptable. This was particularly important for this task, since we had difficulties to normalise the data, as we will describe now.

The balanced results table contained 2112 observations. We generated a boxplot (Figure 16) and a histogram for reading times (TR_reading), our second measure to remove outliers (Figure 17). The boxplot evidenced that there were many outliers, hence the histogram showed that the data did not follow a normal distribution, as we can see in Figures 16 and 17:

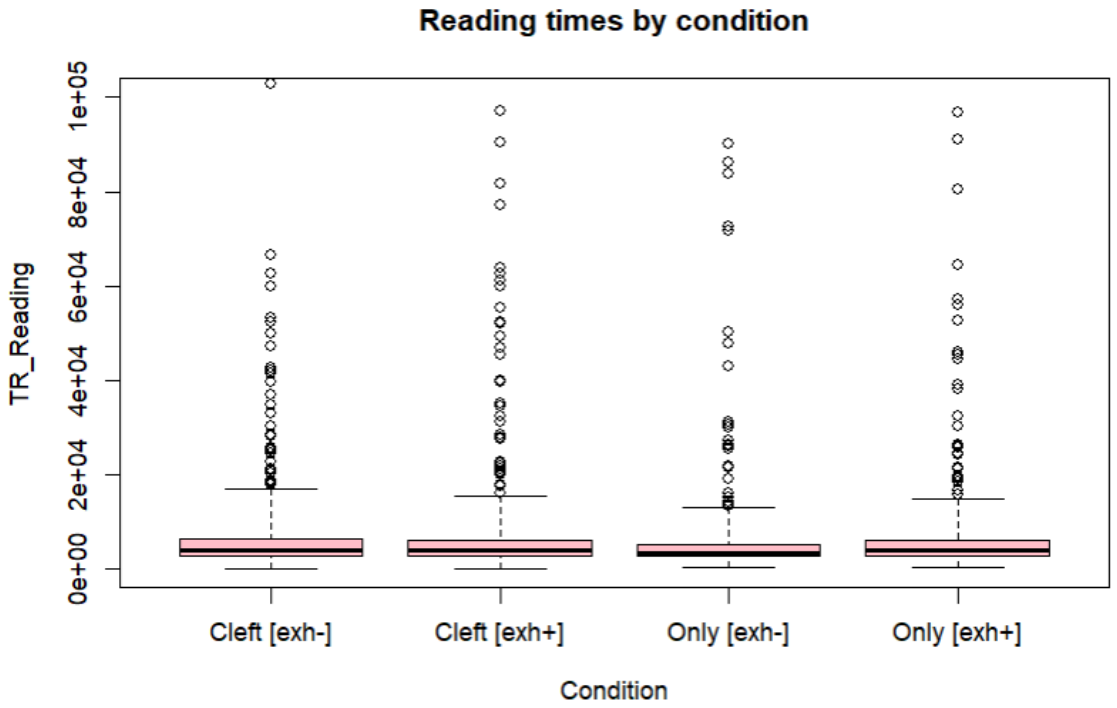


Figure 16 – Boxplot of TR_reading per condition in AJT2.

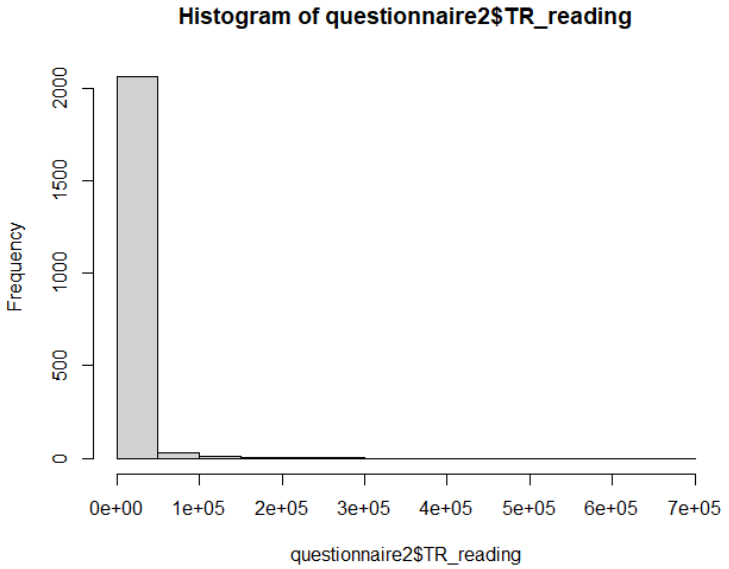


Figure 17 – Histogram of the AJT2 before the application of filters.

After developing a descriptive analysis of the TR_reading measures (Min.: 118; 1st Qu.: 2789; Median: 3881; Mean: 8771; 3rd Qu.: 5994; Max.:679,192), we noticed that the mean was significantly higher than the median, and that the maximum value of the parameter was an extreme outlier. Both of these observations indicate very expressive positive asymmetry, which could be a problem for linear models that assume symmetrical residuals.

To solve this problem, we chose to apply a filter based on percentiles, instead of a numeric filter, as the former are more flexible and adjustable to asymmetric data such as ours. We calculated the upper limit of 90% of all our TR_reading values – which was 10,991.8ms – and then we deleted all the values that corresponded to the remaining 10% (over 10,991.8ms). This left us with 1732 observations on the table, but still did not normalize data, as we noticed by running a *Lilliefors (Kolmogorov-Smirnof)* normality test (D = 0.043424, p-value = 1.402e-08). As a second alternative, we used numeric filters (915ms < TR_reading < 7935ms), which left us with a table of 1740 observations (8 observations more than the previous filter). These filters did not normalise data either, but had a better performance on this purpose, as results from *Lilliefors (Kolmogorov-Smirnof)* were closer to normality (D = 0.02564, p-value = 0.009681).

It is important to consider that: 1) we were facing difficulties to normalise the data; but the extreme outliers had already been removed; 2) we did not want to remove a larger number of observations, as this could impair our statistical analyses; 3) we have already used the ungrammatical sentences as parameters to control participant's attention; 4) it is common for large samples and for reading times not to follow a normal distribution. For all of these reasons, we decided not to keep trying to normalise data. Instead, we checked the residuals' distribution when we dealt with TR_reading, the first DV that we analysed.

We modeled a LMM (*lmer*) to check if there was a difference between the times of reading (TR_reading) across conditions. Random intercepts were included for both participants and items. The LMM revealed considerable variability between participants (n = 44; SD = 0.03), while item-level variability was smaller (n= 192; SD = 0.01).

The non-exhaustive condition for the 'only'-sentences had the shortest mean time of reading, in comparison to the other conditions (Figure 18), and this is consistent with what we found in task 1.

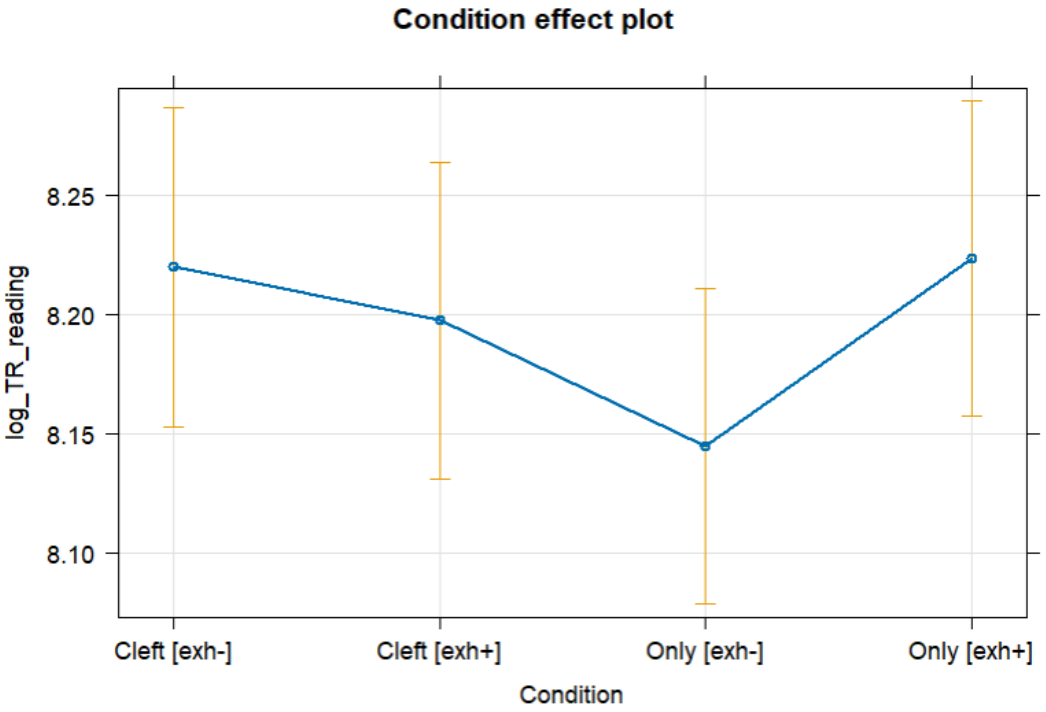


Figure 18 – Log reading times across conditions in AJT2.

The model (Table 6) revealed a similar behaviour between AJT1 and AJT2: a significant effect of exhaustivity that depended on the structure. Variance was estimated for participants ($\tau_{00} = 0.03$) and items ($\tau_{00} = 0.01$):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.20	8.13 – 8.26	**<0.001**
Structure [Only]	0.03	-0.03 – 0.08	0.344
Exhaustivity [Non_exhaustive]	0.02	-0.03 – 0.08	0.426
Structure [Only] × Exhaustivity [Non_exhaustive]	-0.10	-0.18 – -0.02	**0.010**
Random Effects			
σ^2	0.12		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.03		
ICC	0.24		
$N_{\ Participant}$	44		
$N_{\ Item}$	192		
Observations	1740		
Marginal R^2 / Conditional R^2	0.006 / 0.244		

Table 6 – LMM for reading times in AJT2.

These findings (Table 7) match the results of the previous task, when non-exhaustive 'only'-sentences were also significantly faster than exhaustive clefts:

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.22	8.15 – 8.29	**<0.001**
Condition [Cleft [exh+]]	-0.02	-0.08 – 0.03	0.426
Condition [Only [exh-]]	-0.08	-0.13 – -0.02	**0.007**
Condition [Only [exh+]]	0.00	-0.05 – 0.06	0.892
Random Effects			
σ^2	0.12		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.03		
ICC	0.24		
$N_{Participant}$	44		
N_{Item}	192		
Observations	1740		
Marginal R^2 / Conditional R^2	0.006 / 0.244		

Table 7 – LMM for reading times comparing conditions in AJT2.

As we were not dealing with a normal distribution of data, we checked the residuals' distribution of the *lmer*. We plotted a Quantil-Quantil (Q-Q) graph that compared the quantiles of our dataset (in an ex-Gaussian distribution) with the quantiles of a Gaussian (normal) distribution. In Figure 19, the red line represents a normal distribution, while the dots along it correspond to our data. As we can see, there is little or no deviation from our data from the red line, which means that the residuals are not problematic to our model.

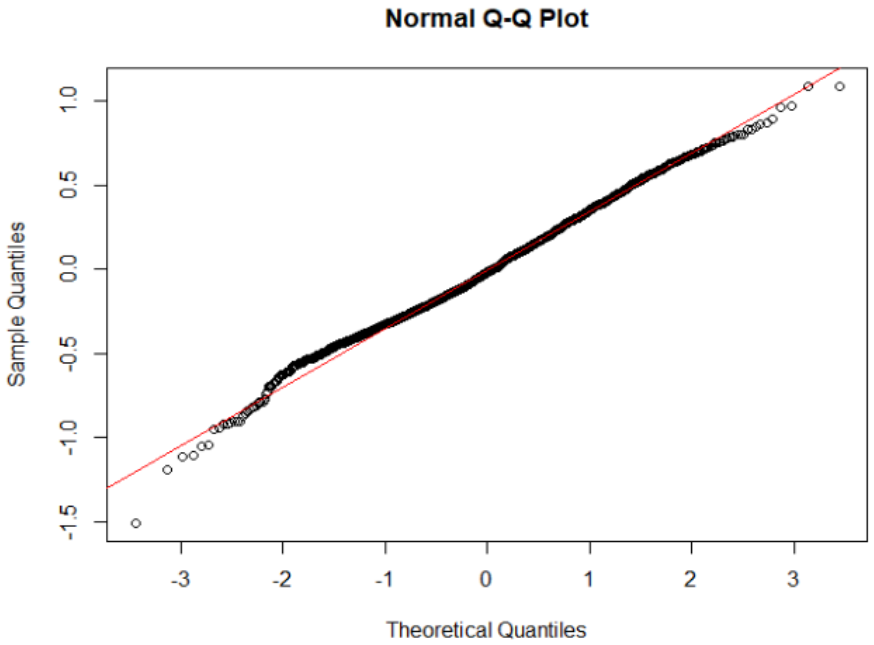


Figure 19 – Normal Q-Q plot of the residuals of AJT2.

The other graph that we plotted to check residuals was a plot of standardized residuals versus fitted values. The resulting plot displayed a dense, symmetrical cloud of points scattered around the horizontal line at zero, with no visible patterns, curves, or funnel-shaped distributions. This indicates that the residuals were homoscedastic and randomly distributed, suggesting that the model captured the structure of the data appropriately (Garcia, 2021). The lack of systematic deviations supports the assumption of linearity and reinforces the validity of the model (Figure 20).

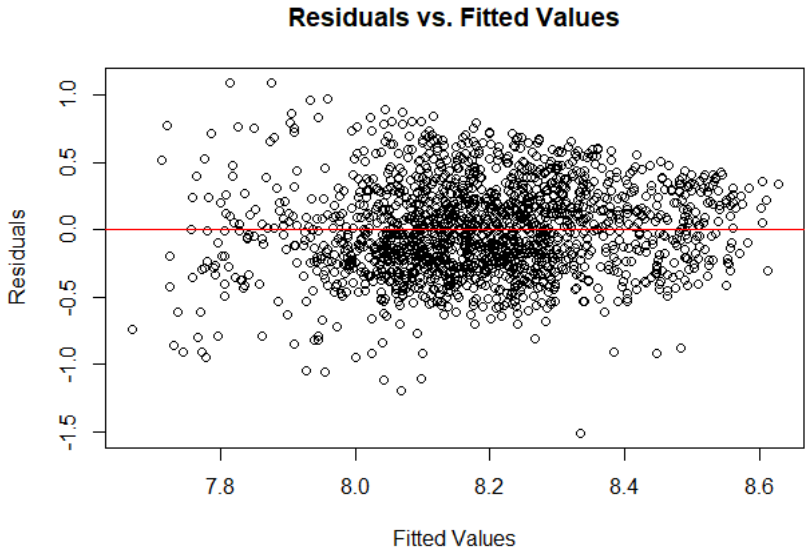


Figure 20 – Residuals vs. fitted values in AJT2.

It is important to note that the differences concerning the length of the stimuli were more subtle in AJT2 than in AJT1. Let us take for example the set of conditions below (10):

(10) a. Yesterday, **only Alice** made coffee and also tea. ([+exh])
b. Yesterday, **only Alice** made coffee and also **Lidia**. ([-exh])
c. Yesterday, **it was Alice who** made coffee and also tea. ([+exh])
d. Yesterday, **it was Alice who** made coffee and also **Lidia**. ([-exh])

Exhaustive clefts (10c) were longer than exhaustive 'only'-sentences (10a), and non-exhaustive clefts (10d) were longer than non-exhaustive 'only'-sentences (10b). Like in the stimuli of AJT1, that was because of the largest number of lexical items ('*it was*/*is*' + '*who*') that it-clefts demanded, in comparison to 'only'-sentences. However, comparing exhaustive and non-exhaustive sentences within the same structure, non-exhaustive clefts (10d) were merely a little longer than exhaustive clefts (10c), and non-exhaustive 'only'-sentences were a little longer than their exhaustive counterparts. That is because we opted in this experiment for the *gapping* ('and also…'), so there were no auxiliary verbs ('*did*'/'*did*') included in exhaustive conditions, which made them slightly shorter than the exhaustive conditions in task 1.

We compared clefts, and we found no difference for reading times between exhaustive and non-exhaustive clefts (Table 8):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.23	8.16 – 8.29	<**0.001**
Condition [Cleft [exh+]]	-0.02	-0.08 – 0.03	0.404
Random Effects			
σ^2	0.12		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.03		
ICC	0.20		
$N_{\ Participant}$	44		
$N_{\ Item}$	96		
Observations	846		
Marginal R^2 / Conditional R^2	0.001 / 0.205		

Table 8 – LMM for reading times for it-clefts in AJT2.

This was the same result that we found in AJT1. Differently to what happened with clefts, exhaustivity for 'only'-sentences increased participants' reading times – as shown in Table 9, the intercept is the non-exhaustive 'only' condition:

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.14	8.07 – 8.21	**<0.001**
Condition [Only [exh+]]	0.08	0.03 – 0.13	**0.004**
Random Effects			
σ^2	0.12		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.04		
ICC	0.27		
$N_{\ Participant}$	44		
$N_{\ Item}$	96		
Observations	894		
Marginal R^2 / Conditional R^2	0.010 / 0.275		

Table 9 – LMM for reading times for 'only'-sentences in AJT2.

We observed the same interesting effect for AJT1: non-exhaustive ‘only’-sentences had the lowest reading times mean, even though they were longer than exhaustive ‘only’-sentences.

After having observed what happened to the reading time across conditions, we started a descriptive analysis of the acceptability rates. Figure 21 illustrates the overall distribution of proportions:

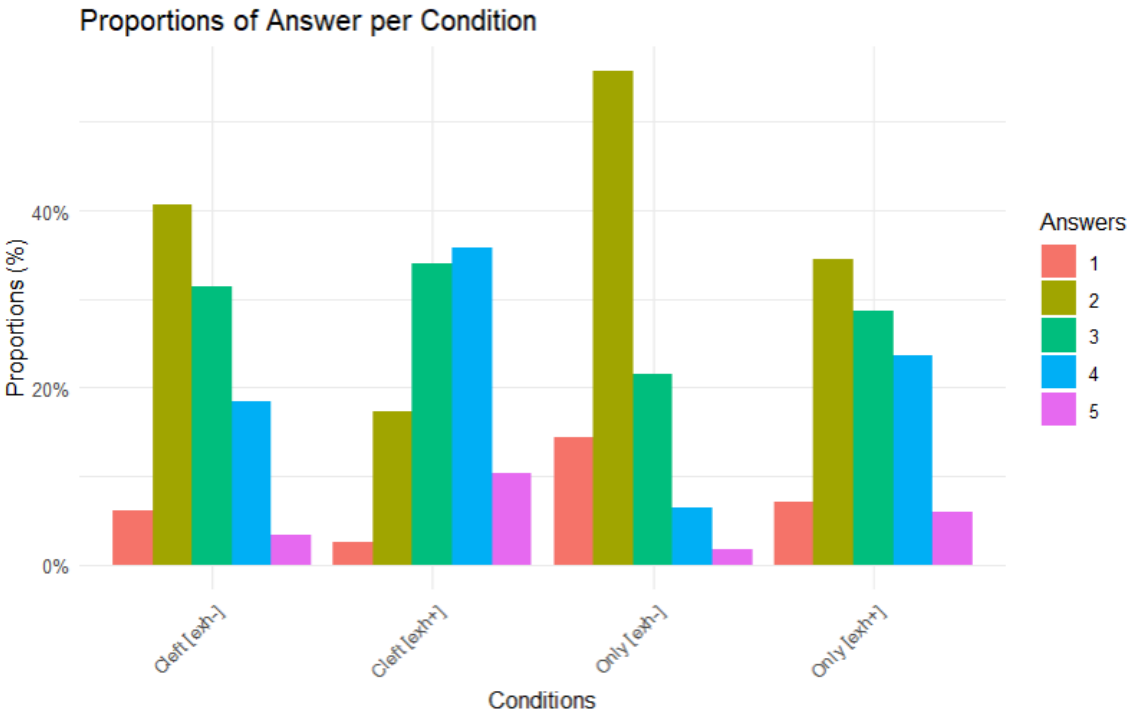


Figure 21 – Acceptability rating across conditions in AJT2.

As we expected, the non-exhaustive condition for 'only'-sentences had the lowest acceptability rate, which is evidenced by the highest percentages of answers 1 and 2 (70%), versus 21.6% and 8.4 for answers 3 and 4 and 5, respectively. The exhaustive condition of the clefts was the most frequently rated as acceptable (answers 4 and 5), with a 46.2% rate, versus 34% neutral (3) and 19.8% unacceptable (1 and 2). None of these overall results differs from our findings in AJT1.

It is worth noting that exhaustive conditions did not violate exhaustivity, while non-exhaustive conditions did. We were, once again, surprised with acceptability rates for the exhaustive 'only'-sentences, which turned out to be considered unacceptable for the most part of participants (unacceptable = 41.65%; neutral = 28.73%; acceptable = 29.62%).

In AJT2, people rated non-exhaustive clefts mostly as neutral (unacceptable = 35%; neutral = 40%; acceptable = 25%). Contrastively, non-exhaustive clefts seemed to be a little more unacceptable (unacceptable = 46.73%; neutral = 31.48%; acceptable = 21.79%) than before (in AJT1) than the clefts from AJT1. This could be because the *gapping* of the 'and also…' structure (AJT2) made the clefts more complex to process than the VP-ellipsis of the AJT1. This probably contributed to a more flexible judgement in AJT1 by the participants,

whereas more complex sentences – like the ones in AJT2 – might have influenced their decision negatively.

Our descriptive findings showed that: 1) non-exhaustive 'only'-sentences had faster reading times, and the lowest acceptability rate; 2) non-exhaustive sentences (clefts and 'only') were more often rated as unacceptable; 3) exhaustive 'only'-sentences had a higher percentage of unacceptability over exhaustive clefts.

An exploratory Pearson's chi-squared test revealed a significant association between condition and response type ($\chi^2(2) = 92.32$, $p < 0.001$). This suggests that the distribution of responses varied across experimental conditions (Figure 22):

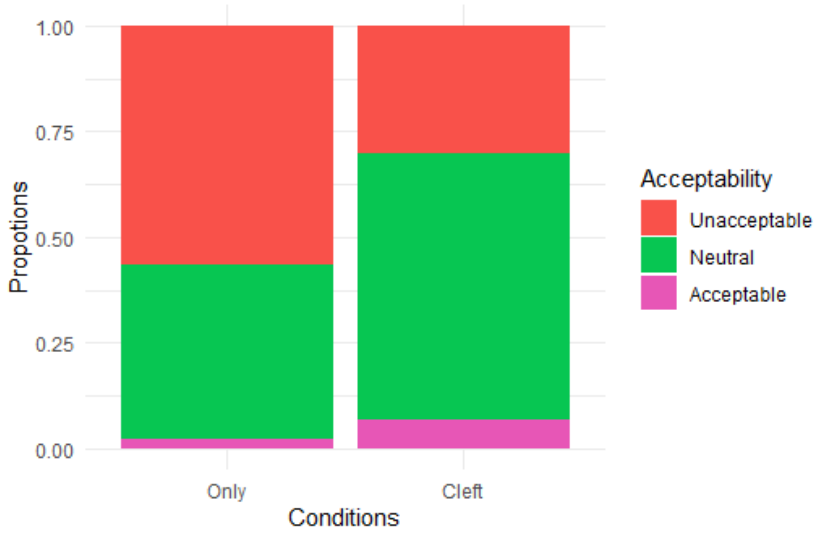


Figure 22 – Proportions of acceptability ratings across syntactic structure in AJT2.

Subsequently, we checked the association between exhaustivity and acceptability ($\chi^2(2) = 144.36$, $p < 0.001$), which seemed to be stronger than syntactic structures (Figure 23):

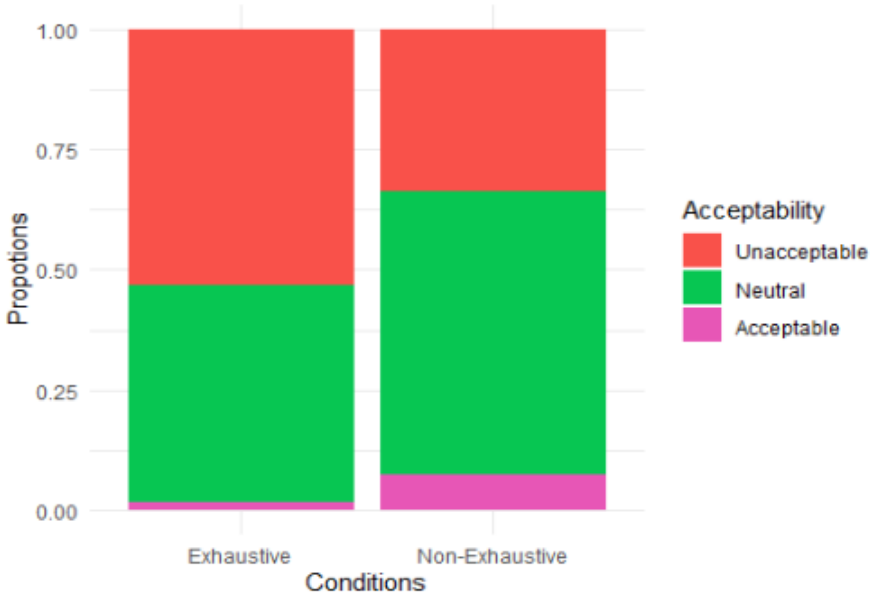


Figure 23 – Proportions of acceptability ratings across exhaustivity conditions in AJT2.

We used a CLMM model, similar to the one that we used to analyse data from AJT1: this model estimates the cumulative log-odds of a higher response on the acceptability scale (DVs), in relation to the interaction between syntactic structure and the exhaustivity (IVs). The DV variable was treated as an ordered factor ranging from 1 (extremely unacceptable) to 5 (extremely acceptable), and the model included participants and items as random effects.

Differently from what we observed for AJT1, the interaction between syntactic conditions and exhaustivity was not significant in this task (Table 9). Therefore, 'and also…' structures present additive effects of structure and exhaustivity rather than an interactive effect. This means that 'only'-sentences generally reduced the chances for an 'acceptable' rating (OR = 0.35 , 95% CI [0.25, 0.47], p = < 0.001), and that non-exhaustive decreased the chances of acceptability (OR = 0.26 , 95% CI [0.19, 0.35], p < 0.001), but the impact of one factor did not depend significantly on the other (OR = 0.85, 95% CI [0.54, 1.32], p = 0.466).

	Answer		
Predictors	*Odds Ratios*	*CI*	*p*
1|2	0.01	0.01 – 0.01	**<0.001**
2|3	0.20	0.13 – 0.30	**<0.001**
3|4	1.20	0.80 – 1.81	0.381
4|5	12.07	7.73 – 18.85	**<0.001**
Structure [Only]	0.35	0.25 – 0.47	**<0.001**
Exhaustivity [Non_exhaustive]	0.26	0.19 – 0.35	**<0.001**
Structure [Only] × Exhaustivity [Non_exhaustive]	0.85	0.54 – 1.32	0.466
Random Effects			
σ^2	3.29		
$\tau_{00\ Item}$	0.21		
$\tau_{00\ Participant}$	1.30		
ICC	0.31		
$N_{\ Participant}$	44		
$N_{\ Item}$	192		
Observations	1740		
Marginal R^2 / Conditional R^2	0.152 / 0.419		

Table 10 – CLMM with interaction of the variables in AJT 2.

The conclusions for the acceptability task 2 were: 1) 'only' was associated with a significantly lower likelihood of receiving higher rates of acceptability; 2) non-exhaustive conditions decreased acceptability, in relation to exhaustive conditions, and this effect did not depend on the syntactic structure; 3) there was no significant effect from the interaction of the IVs.

4.1.2.7. Partial Conclusions – AJTs 1 and 2 (in English)

Different analyses were conducted with two kinds of variables of these tasks: reading times (TR_reading) and answers (1-5). Descriptive analyses presented the reading times

means and raw proportions of responses by condition, providing an initial overview of the data distribution. There were also exploratory chi-square tests to analyse the distribution of acceptability across conditions. However, these summaries do not account for variability across participants and items, nor do they capture the combined effects of experimental predictors.

To address these limitations, the main analyses used mixed-effects models. Linear mixed models (LMMs) were used for reading times, and cumulative link mixed models (CLMMs) for acceptability ratings. These models included random intercepts for participants and items, thereby accounting for individual differences and item-specific variability, while allowing for the simultaneous evaluation of syntactic structure, exhaustivity, and their interaction. Importantly, the CLMM respects the ordinal nature of Likert-scale responses (Fonseca; Maia, 2022) by modeling cumulative probabilities across ordered categories (Joshi *et al.*, 2015). After analysing the acceptability experiments (AJTs) conducted in English, we noticed some patterns.

Firstly, both AJTs 1 and 2 revealed significantly faster reading times for non-exhaustive ‘only’-sentences, in comparison to non-exhaustive clefts and to exhaustive ‘only’-sentences. That is because the interaction between ‘only’ and non-exhaustivity produced a main effect of decrease in reading times. As discussed in previous sections, reading times could be considered to contemplate the actual judgement of the sentence. Thus, such effects might suggest that ‘only’-sentences with exhaustivity violations were judged faster than any other sentences.

Secondly, we noticed main effects of structure and/or exhaustivity, with ‘only’ and non-exhaustivity decreasing the odds of a sentence to be rated as acceptable. In other words, the adverb ‘only’ and the violation of the semantic exhaustivity contribute significantly for a sentence to be considered unacceptable. However, results show that such independent variables had an effect when they were isolated or combined in AJT1 (‘... did too’), which differs from AJT2 (‘and also…’), as the interaction between structure and exhaustivity produced no significant effect in the latter. Put differently, in AJT2 the combination of being an ‘only’-sentence and being non-exhaustive attenuated the effect that each variable produced when isolated.

One possible explanation for this difference lies in the syntactic realization of the additive continuation. In AJT1, violations were constructed using VP-ellipsis (‘…did too’),

which explicitly reintroduces the same event with a different subject (Frazier; Clifton, 2001). This makes the addition salient and conflicts directly with the semantic effect of exhaustivity. On the other hand, AJT2's use of 'and also…' involves a *gapping* structure (Carlson, 2002), which may allow for a more flexible or less explicit interpretation of the added information. As a result, the additive completion gathered with a contrastive particle ('only') may be less strongly perceived as a violation, reducing the likelihood of an interaction effect.

4.1.3. Acceptability Task 3 – AJT3 in BP

Analogously to the acceptability tasks 1 and 2 in English, we developed a Likert acceptability test in BP. For this, we used the same design and type of items that we have previously described for the experiments in English.

4.1.3.1. Materials

All of the experimental sentences in PB had the following structure: temporal adverbial + ('*só*'/*clivagem*) subject-NP + transitive verb + '*e também* (subject/object constituent)'[20]. We have to take into account the differences of language, and that is why we dedicate this section to expose the reasons for our choices and understandings of each structure.

We have presented – in section 3.2 – studies that argue that '*só*' in BP shares syntactic, semantic and pragmatic similarities to 'only' in English (Lima, 2010; Müller; Martins, 2006; Lignani; Fonseca, 2023). These studies show that '*só*', like 'only', associates with focus and exhaustive interpretations.

Likewise, the construction '*e também*' shares similar features to 'and also'. The conjunction '*e*' is the Brazilian Portuguese counterpart of the coordinate conjunction 'and' (Krifka, 1999; Camacho, 2000). Both '*também*' and 'also' are additive particles, hence essentially incompatible with focus particles that convey exclusiveness, such as '*só*' and 'only' (König, 1991; Kiss, 1998; Zimmerman, 2020; Berger; Höhle *et al*., 2009).

In this specific experimental context, the BP construction '*e também…*' is more similar to 'and also…' (used in AJT2), in comparison to 'did…too' (used in AJT1). The

[20] This pattern would translate as temporal adverbial + ('only'/it-cleft) subject-NP + transitive verb + 'and also (subject/object constituent)', which is exactly equivalent to the pattern that we used in the acceptability task 2 in English.

reason for that is simple: '*e também*...' equals 'and also…' in terms of syntactic and semantics because both occur preverbally[21], and they differ from '…did too', where the additive particle ('too') occurs post-verbally.

Additive particles might often be incompatible to *clivadas totais*/it-clefts, if we admit that the pattern of such strategies is to convey exhaustivity (Horn, 1981; Kiss, 1998; Krifka, 2008; Drenhaus *et al*., 2010). Concerning clefts, we have addressed them in section 3.2. We evidenced that BP's *clivadas totais* equal English *it-clefs*, in terms of structure (syntax) and semantics (Reis, 2005; Braga; De Mello Barbosa, 2009; Kato, 2013; Quarezemin; Tescari Neto, 2015). That is the reason why we chose to use *clivadas totais* in this task, among other kinds of clefts of the BP (Braga; De Mello Barbosa, 2009; Alves *et al*., 2015).

The stimuli in the task in BP were 24 (twenty-four)[22] sets of sentences in four different conditions, 28 (twenty-eight) fillers and 3 (three) practice-sentences. The 28 fillers were divided in 14 grammatical and 14 ungrammatical[23]), and 3 sentences of practice, displayed one by one before the actual trial. Half of the sets (12) presented female subject-NPs and the other half of them presented male ones, with the purpose to minimise gender biases. In addition, we made sure that half of the subject-NPs (12) were common nouns (*substantivos comuns*) and the other half were proper nouns (*substantivos próprios*) (Huddleston; Pullum, 2002; Bechara, 2018).

Each set contained a sentence that varied across 4 (four) different conditions: a) with the adverb '*só*', conveying exhaustivity ([+exh]); b) with the adverb '*só*', but non-exhaustive ([-exh]); c) *clivada total*, conveying exhaustivity ([+exh]); and d) *clivada total*, but non-exhaustive ([-exh]).

For example, the set number (11) below:

(11) a. No shopping, **só** a Jéssica comprou roupas e também sapatos. ([+exh])

At the mall, only Jessica bought clothes and also shoes.

b. No shopping, **foi** a Jéssica **quem** comprou roupas e também a Nicolle. ([-exh])

At the mall, only Jessica bought clothes and also Nicolle.

---

[21] As we have discussed in subsection 4.1.2 Acceptability task 2, '*e também*…' and 'and also…' would be syntactic-semantic counterparts because they demand a *gapping* process, since they are preverbal. This is different from the VP-ellipses that happen in the structure '…did too' (Johnson, 2009).

[22] See Appendix C for all the items.

[23] We included ungrammatical items, such as '**Ladrões atacaram a Helena e com fugiram polícia*' ('*Thieves attacked Helena and with ran way Police'), in this experiment for the same reasons we did it with the tasks in English. The first one was to conceal experimental items in a more effective way, and the second one to test participants' attention during the trial.

c. No shopping, **foi** a Jéssica **quem** comprou roupas e também sapatos. ([+exh])
At the mall, it was Jessica who bought clothes and also shoes.
d. No shopping, **foi** a Jéssica **quem** comprou roupas e também a Nicolle. ([-exh])
At the mall, it was Jessica who bought clothes and also Nicolle.

In this example, the object-NP *'shoes'* preserves the exhaustivity expected by '*only*', because it did not conflict with the idea that Jessica is the one person who buys both clothes and shoes. Therefore, sentences (11a) and (11c) are exhaustive.

On the other hand, adding 'Nicolle' as a subject-NP for the same action ('buying clothes') represents a disruption to the idea that there was exclusively one person buying clothes, who would be Jessica. Consequently, sentences (11b) and (11d) are non-exhaustive, as they present a violation of the exhaustivity supposedly predicted until people read a second subject-NP.

4.1.3.1 Method

This experiment, as the experiments in English, was designed on the website PC IBEX (Zehr; Schwarz, 2018), and shared with participants who applied on the SONA SYSTEMS (2024). Again, people could execute this task remotely. We built the trial using the Latin square design (Abbuhl; Gass; Mackey, 2013), with four versions of the experiment (A, B, C and D). We included on the PC IBEX script an automatic randomization for items and versions of the experiment. Once participants accessed the link provided, they saw an initial page (Figure 24)[24]:

[24] Figure 24 shows the information requested for the participants in the task in PB: a) full name ('*nome completo*'); b) e-mail address; c) age ('*idade*'); and d) education level ('*escolaridade*'). Note that we did not request participants' names or e-mails for the acceptability tasks 1 and 2 (in English). That was because the people who participated in the tasks in English were all students of Psychology in the University of Leicester, and the SONA Systems did not allow us to collect any personal data. In those tasks, participants' identities remained anonymous from their first access to the tasks' links, and their results (answers, measures, etc) were displayed on the results' table under a SONA ID encoded number. This process guaranteed anonymity and credits for participants. Differently, for the BP task, we had to ask participants' names and valid e-mail addresses so we could send them their Participation Certificate, as we did not have a system, like SONA, to do that automatically.

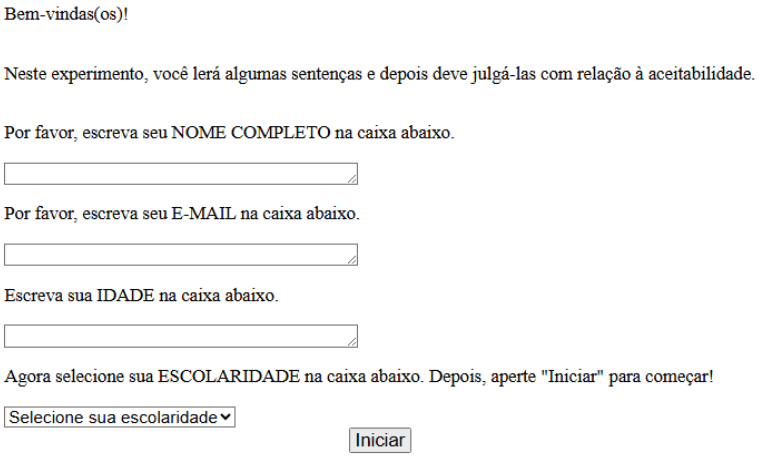


Figure 24 – Welcome screen of the AJT3..

After filling the forms with the information requested, participants read the instructions on the next page (Figure 25), that were the same instructions of the tasks in English:

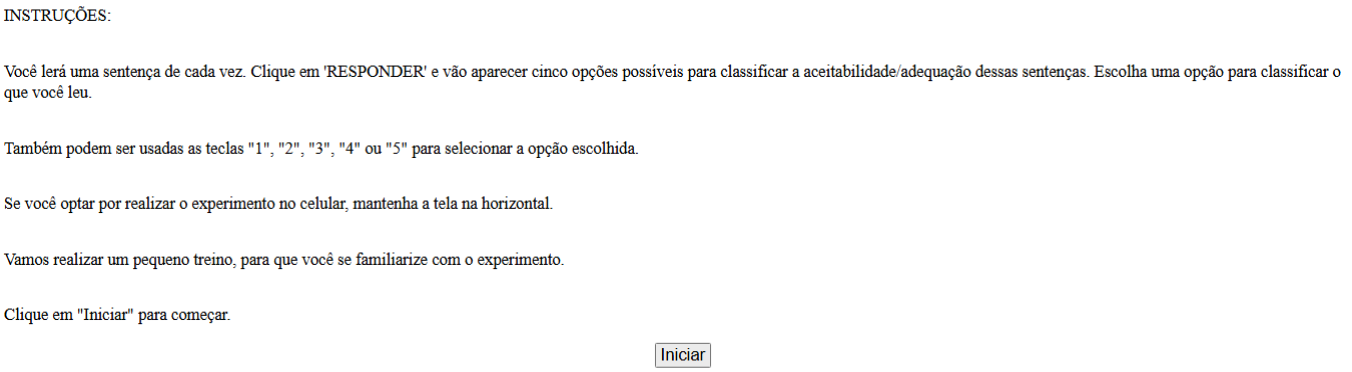


Figure 25 – Instructions of the AJT3.

Like for the tasks in English, there were 3 sentences before the experimental items, that displayed on the screen until participants pressed the '*Responder*' ('*Answer*') button (Figure 26):

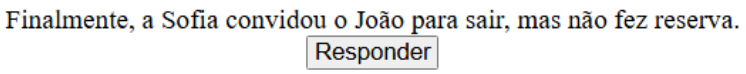


Figure 26 – Example of practice-sentences of the AJT3.

After having read a sentence and clicked the button '*Responder*' (Figure 26), participants should answer the acceptability question '*Quão aceitável essa frase parece para você?*' ('How acceptable is this phrase for you?') (Figure 27). As with the tasks in English, the acceptability scale varied from 1 to 5, being 1 the least acceptable.

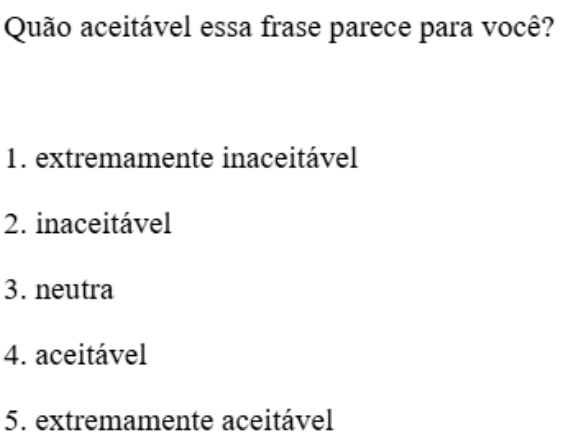


Figure 27 – Acceptability scale of the task in BP.

Participants could answer the question by pressing the correspondent keys on their keyboards (1, 2, 3, 4 or 5) or by selecting their answer with their mouse. Once they completed the 3-sentence practice, they were ready to begin the experiment (Figure 28):

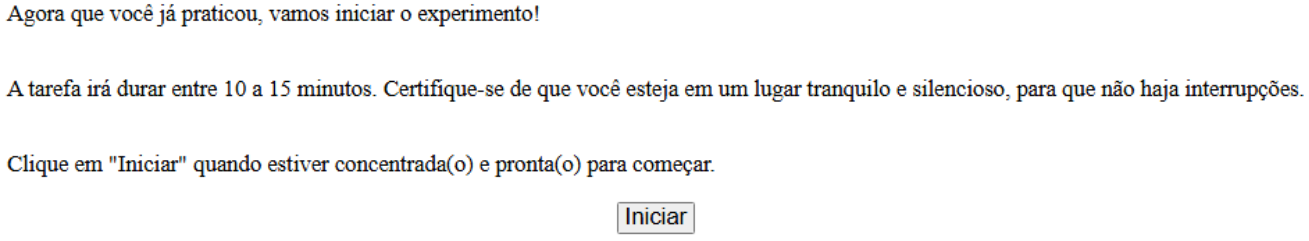


Figure 28 – Beginning of the experiment of the AJT3.

Figure 29[25], shows an item (randomised) and displayed on the screen:

[25] Figure 29 shows the item '*Durante a aula, só a Beatriz termina os exercícios e também a Renata*, which translates as 'During the class, only Beatrice finishes the exercises and also Renata'.

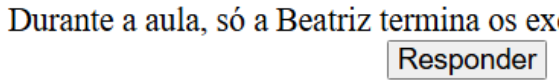


Figure 29 – Example of an experimental item of the AJT3..

After the participants have gone through the 55 sentences (24 experimental ones, 28 fillers, plus 3 practice-sentences), a message appeared on the last screen of the trial. On the last screen (Figure 30), there was a link to Google Forms, where participants' could read the terms of the research and consent about usage of data produced by them throughout the trial. After they gave their consent, we sent a Certificate of Participation to the e-mails that they have provided by filling the initial screen.

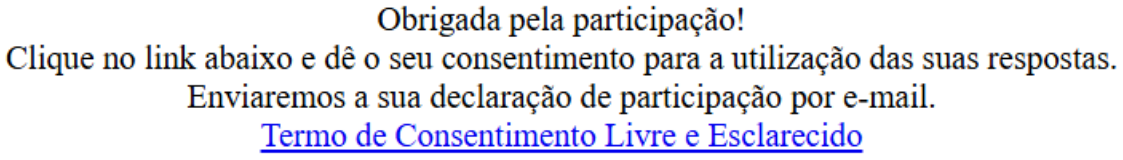


Figure 30 – Final screen of the AJT3.

4.1.3.2 Independent Variables

The independent variables for this task in BP were the same for the tasks in English: exhaustivity and the syntactic structure of the items. Exhaustivity defined exhaustive and non-exhaustive conditions, while the syntactic structure varied in '*só*'-sentences and *clivadas*.

4.1.3.3 Dependent Variables

The dependent variables (DVs) of this experiment were the different marks on a 1-to-5-scale of acceptability (Figure 27) and the reading times (TR_reading).

4.1.3.4 Participants

Forty-five (45) native speakers of BP participated in this experiment. Before initiating the data analysis, we balanced the table of results and remained with data from 40 people, 10 participants for each version of the experiment (A, B, C or D). The mean age was 27.9 years

(min. 18; max. 36), 35 women and 5 men. Among participants, we had 17 undergraduates, 8 graduated and 15 postgraduates.

Everyone who participated in this experiment did so voluntarily. We shared the link to the experiment on social media, especially among the academic community and those who wanted to be volunteers for this research could take the experiment remotely. As shown in Figure 30, at the end of the experiment, participants had access to a link that led them to a Consent Form (‘*Termo de Consentimento Livre e Esclarecido*’– TCLE), where we describe the task in general. This Consent Form (TCLE) made clear for the volunteers that their participation was free of charges or compensations, and that all the data provided by them would remain completely anonymous. After giving their consent, participants received an e-mail with a Declaration of Participation on the study, as a PDF file.

4.1.3.5 Summary for the AJTs 1 and 2

The data of this experiment was analysed on RStudio (R Core Team, 2025). We removed from the results’ table data from participants whose accuracy rate of (un)acceptability for ungrammatical fillers was lower than 80%. The balanced results table contained 960 observations, then we generated a boxplot and a histogram for TR_reading, our secondary filter to remove outliers. It is worth recalling that the TR_reading was the reading time of each sentence, corresponding to the time between items were shown on the screen and the moment when participants pressed the ‘Answer’ button, that would lead to the acceptability scale on the next screen. The boxplot and the histogram showed that we had some outliers and that the data did not follow a normal distribution, as we can see in Figures 31 and 32:

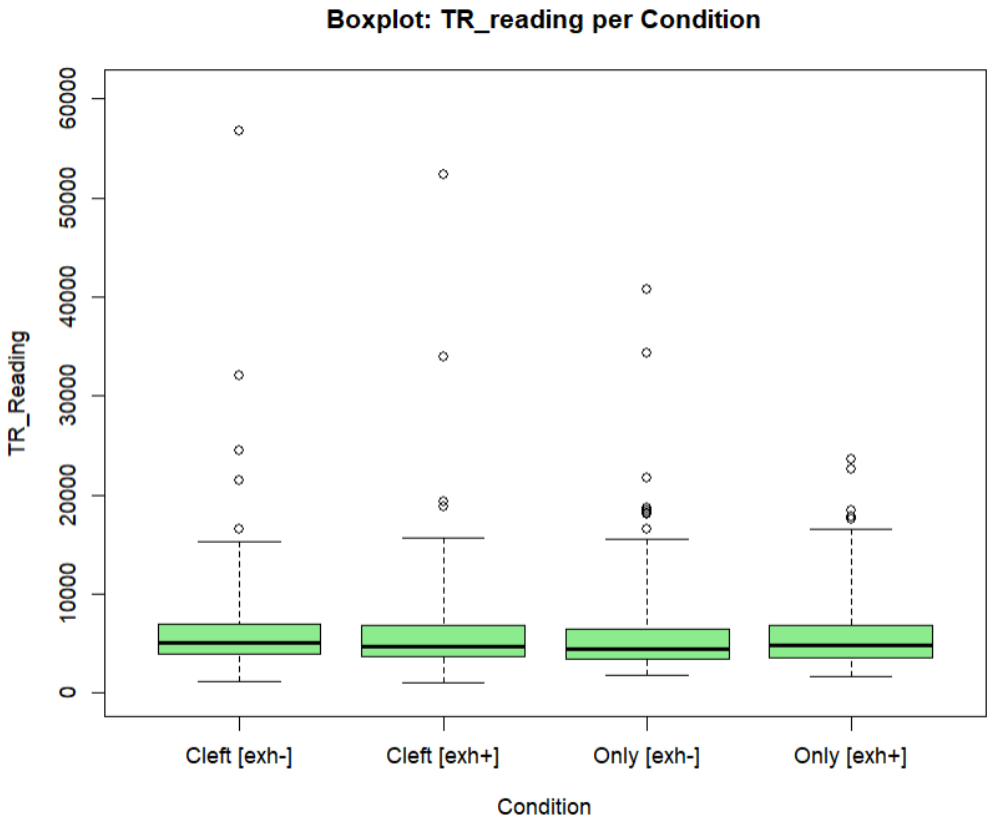


Figure 31 – Boxplot of TR_reading per condition in AJT3.

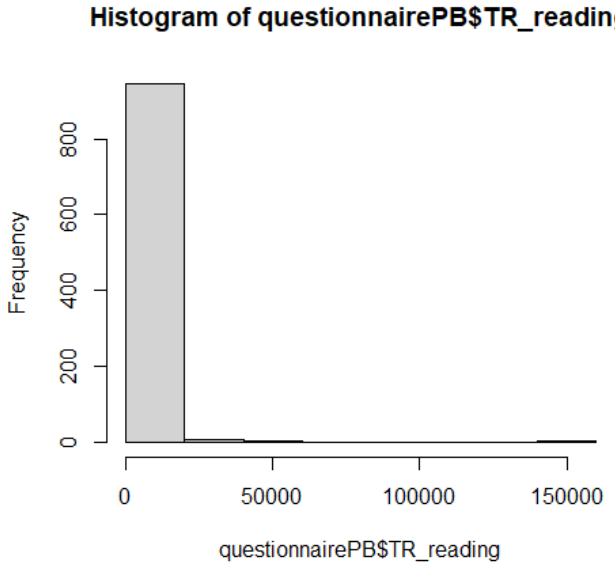


Figure 32 – Histogram of ATJ3 before normalisation.

After having a descriptive analysis of the TR_reading measures (Min.: 1,098ms; 1st Qu.: 3,601; Median: 4,762; Mean: 6,291; 3rd Qu.:6754; Max.:15,5596), we applied a filter of

TR_reading < 11.600ms. Then, we deleted the outliers and we ran a Lilliefors (Kolmogorov-Smirnof) normality test, which indicated that we were able to normalise data (Figure 33) by filtering and removing outliers (D = 0.029407, p-value = 0.06446).

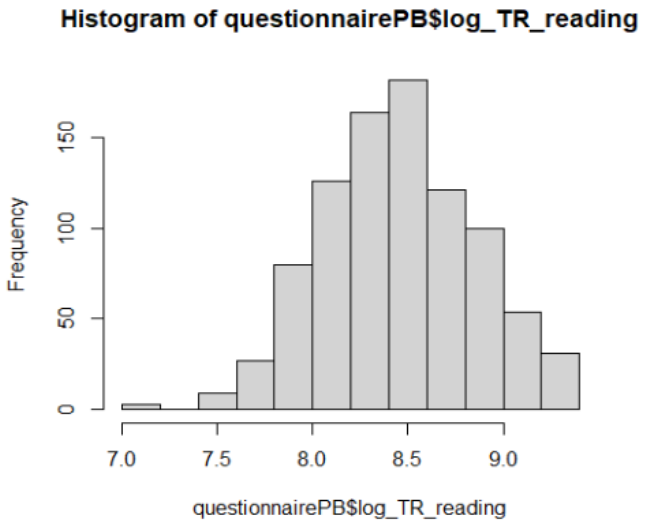


Figure 33 – Histogram of AJT3 after normalisation.

After the removal of the outliers, we analysed data from a table with 897 observations (63 observations less than the original table).

Figure 34 shows the reading times for the task in BP, with a pattern TR_reading that repeats across acceptability tasks:

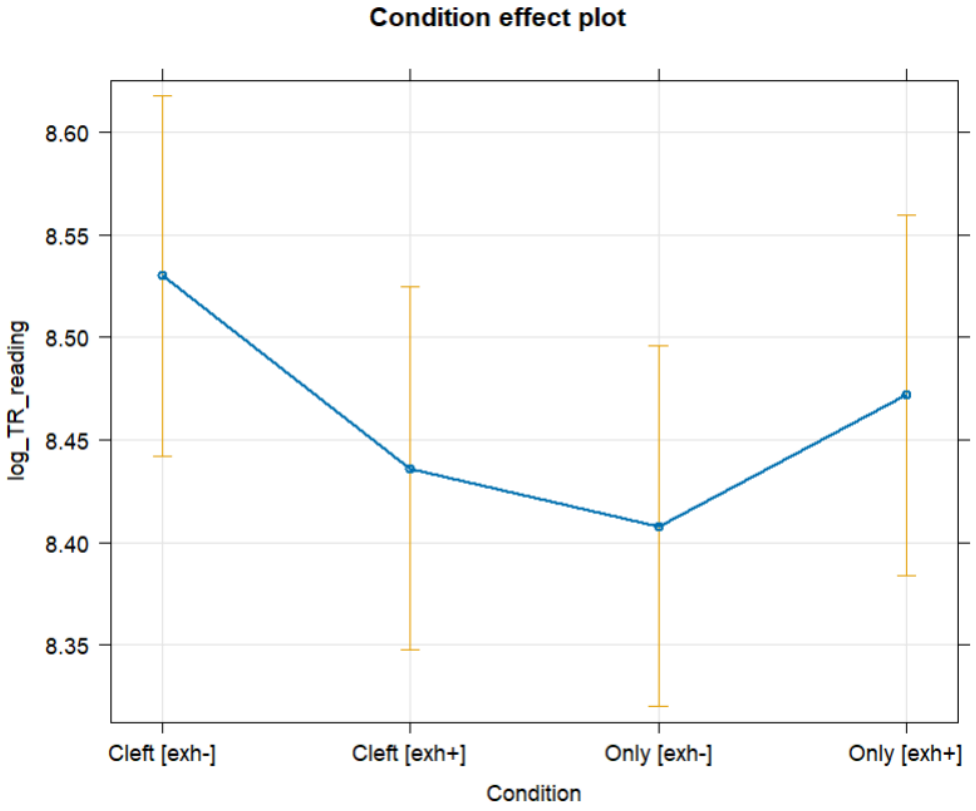


Figure 34 – Log-transformed reading times across conditions in AJT3.

We developed a LMM (*lmer*) to look for effects of structure and/or exhaustivity (TR_reading) across conditions. We included participants and items as random intercepts.

Results (Table 11) show that, differently from what we found for the English tasks, there was a positive effect of non-exhaustivity (β = 0.09, 95% CI [0.01, 0.18], p = 0.027), but not of structure (β = 0.04, 95% CI [-0.05, 0.12], p = 0.401). However, a negative effect of structure appeared when combined to non-exhaustivity (β = -0.16, 95% CI [-0.27, -0.04], p = 0.009). Interestingly, contrasting with AJTs 1 and 2, when the effect of non-exhaustivity decreased reading times, in this AJT3 (in BP), non-exhaustivity produced an increase of reading times.

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.44	8.35 – 8.52	**<0.001**
Structure [Only]	0.04	-0.05 – 0.12	0.401
Exhaustivity [Non_exhaustive]	0.09	0.01 – 0.18	**0.027**
Structure [Only] × Exhaustivity [Non_exhaustive]	-0.16	-0.27 – -0.04	**0.009**
Random Effects			
σ^2	0.11		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.04		
ICC	0.34		
$N_{\ Participant}$	40		
$N_{\ Item}$	96		
Observations	897		
Marginal R^2 / Conditional R^2	0.013 / 0.348		

Table 11 – LMM for reading times in AJT3.

It is important to note that there were differences in the length of the stimuli. Let us take for example the set of conditions below (12):

(12) a. Nas férias, **só o Otávio** jogou <u>baralho</u> e também <u>Uno</u>. ([+exh])
On holidays, only Otavio played cards and also Uno.

b. Nas férias, **só o Otávio** jogou <u>baralho</u> e também **o Rodrigo**. ([-exh])
On holidays, only Otavio played cards and also Rodrigo.

c. Nas férias, **foi o Otávio quem** jogou <u>baralho</u> e também <u>Uno</u>. ([+exh])
On holidays, it was Otavio who played cards and also Uno.

d. Nas férias, **foi o Otávio quem** jogou <u>baralho</u> e também **o Rodrigo**. ([-exh])
On holidays, it was Otavio who played cards and also Rodrigo.

Non-exhaustive clefts (12d) were overall longer than exhaustive clefts (12c). We believe that this might be a natural effect, because exhaustive and non-exhaustive clefts of a given set (like 4) contained the same sequence of lexical items, but the non-exhaustive ones ended in article + NP, as in '*o Rodrigo*' ('Rodrigo')[26]. We remodeled the *lmer* to compare conditions exclusively among the same syntactic structure. Comparing TR_reading times between exhaustive and non-exhaustive clefts, the effect of higher reading times for the latter in comparison to the former were milder (p= 0.035), but still significant (Table 12):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.53	8.44 – 8.62	**<0.001**
Condition [Cleft [exh+]]	-0.09	-0.18 – -0.01	**0.035**
Random Effects			
σ^2	0.11		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.04		
ICC	0.32		
$N_{\ Participant}$	40		
$N_{\ Item}$	48		
Observations	445		
Marginal R^2 / Conditional R^2	0.014 / 0.331		

Table 12 – LMM for reading times for it-clefts in AJT3.

Similarly, we used a LMM comparing exhaustive to non-exhaustive 'only'-sentences to verify if the effect that we have just discovered for clefts could be related to (non-) exhaustivity (Table 13). If we observed a significant effect of reading times between non-exhaustive and exhaustive 'only'-sentences, that would be evidence favouring the

[26] The researchers that designed, applied and analysed these experiments are from Juiz de Fora, MG (Brazil), and so were all the participants that took this task. That means that their variation of BP uses definite articles '*o*'/'*a*' ('the') before proper nouns (Perini, 2010; Schwindt, 2014), and we wanted to preserve this naturalness for the stimuli.

argument that non-exhaustive conditions were more difficult to process than exhaustive ones. This would explain higher TR_reading with the (non-)exhaustivity. Nevertheless, there was no significant effect between 'only'-sentences (p= 0.113):

	log_TR_reading		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.41	8.32 – 8.50	**<0.001**
Condition [Only [exh+]]	0.06	-0.01 – 0.14	0.113
Random Effects			
σ^2	0.10		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participant}$	0.05		
ICC	0.35		
$N_{Participant}$	40		
N_{Item}	48		
Observations	452		
Marginal R^2 / Conditional R^2	0.006 / 0.350		

Table 13 – LMM for reading times for 'only'-sentences in AJT3.

In conclusion, in BP we have no evidence to believe that higher TR_reading means for the non-exhaustive clefts would not be explained by the fact that non-exhaustive clefts (such as 12d) are longer than exhaustive clefts (12c).

It is interesting to notice that the results for this task show an opposite outcome, in comparison to AJT1. For AJT1, we found that the TR_reading was significantly higher in exhaustive 'only'-sentences, when compared to non-exhaustive ones, and that there were no significant effects found for clefts. Differently, this task (AJT3) showed that non-exhaustive clefts increased TR_reading significantly, in comparison to exhaustive clefts, and that there were no significant effects for the conditions with 'only'-sentences.

After having analysed TR_reading as a dependent variable, we continued the analysis of participants' choices. Non-exhaustive 'only'-sentences had the lowest acceptability, which

is evidenced by the highest percentages of answers 1 and 2 (63.3%), meaning that they were the sentences more often rated as unacceptable, versus 17.7% neutral (3) and 20% acceptable (4 and 5). Non-exhaustive clefts, like non-exhaustive 'only'-sentences, contained a violation and were expected to be interpreted as less acceptable than their exhaustive counterparts.

In sum, we would expect non-exhaustive conditions to have the higher rates of unacceptability, which actually appeared to be the case (Figure 35). Besides non-exhaustive 'only'-sentences, non-exhaustive clefts also had unacceptable ratings as a majority answer (38.3%), in contrast to neutral (23.7%) or acceptable (38%) answers. Exhaustive clefts were rated more often as acceptable sentences, with 78.5% selections of 4 and 5, versus 11.5% of 3(neutral) and 10% of 1 and 2 (unacceptable). Exhaustive 'only'-sentences have also shown to be overall acceptable (66%), instead of neutral (14.6%) or unacceptable (19.4%).

These results show that non-exhaustive sentences were more often classified as unacceptable, whereas exhaustive ones were generally rated as acceptable. We did not expect a different result and, differently from what happened with task 1 (in English), we were not surprised. Figure 35 illustrates a higher rate of acceptability for exhaustive conditions:

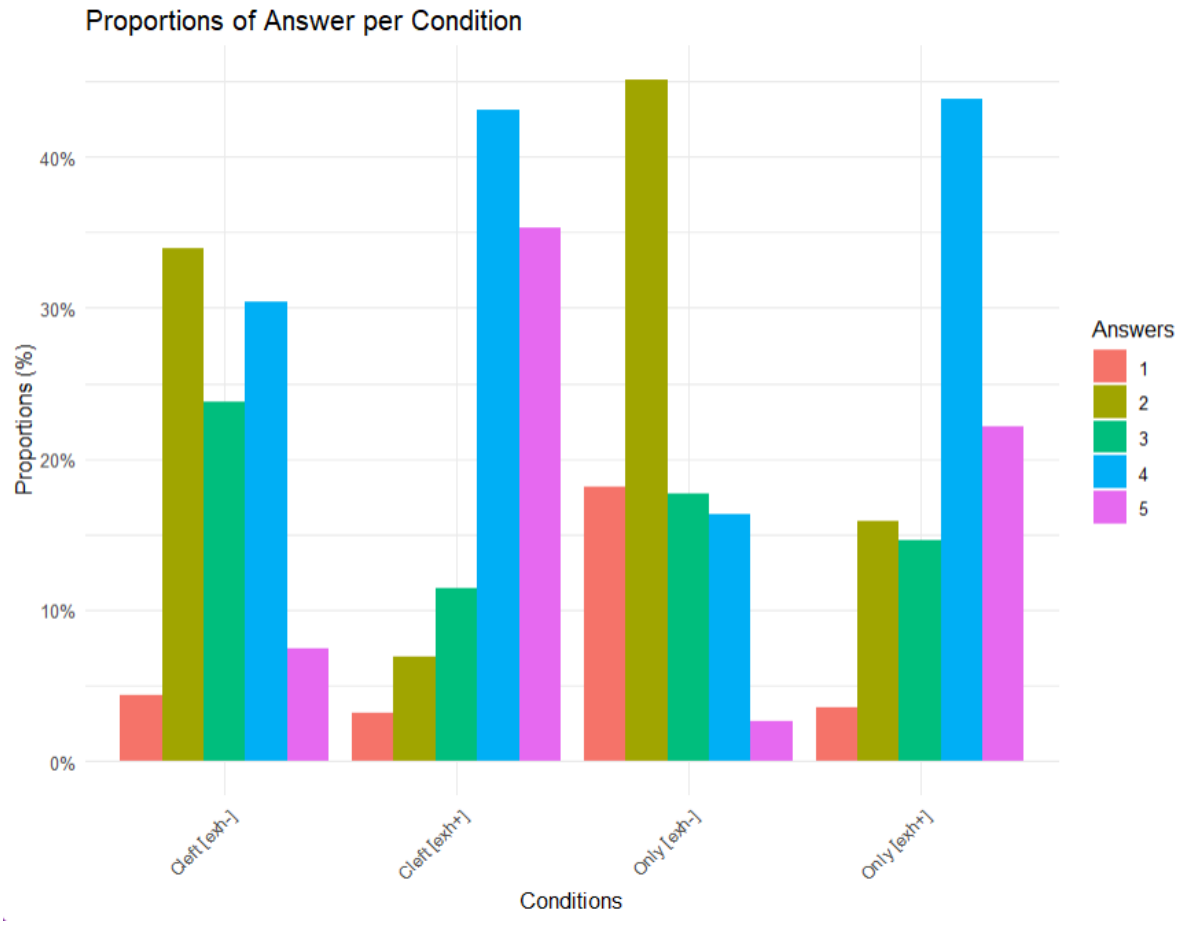


Figure 35 – Acceptability rating across conditions in AJT3.

Our descriptive findings so far show that: 1) non-exhaustive clefts were read significantly slower than exhaustive clefts; 2) non-exhaustive 'only'-sentences with lower acceptability rates were read faster than exhaustive 'only'-sentences; 3) non-exhaustive sentences, regardless of syntactic structure, were less acceptable than exhaustive ones; 4) non-exhaustive 'only'-sentences were more often rated as unacceptable; 5) exhaustive clefts were more often rated as acceptable.

To begin with an inferential analysis, we developed a Pearson's chi-squared test (Figure 36), which indicated an association between syntactic structure and acceptability ($\chi^2(2) = 32.543$, $p < 0.001$):

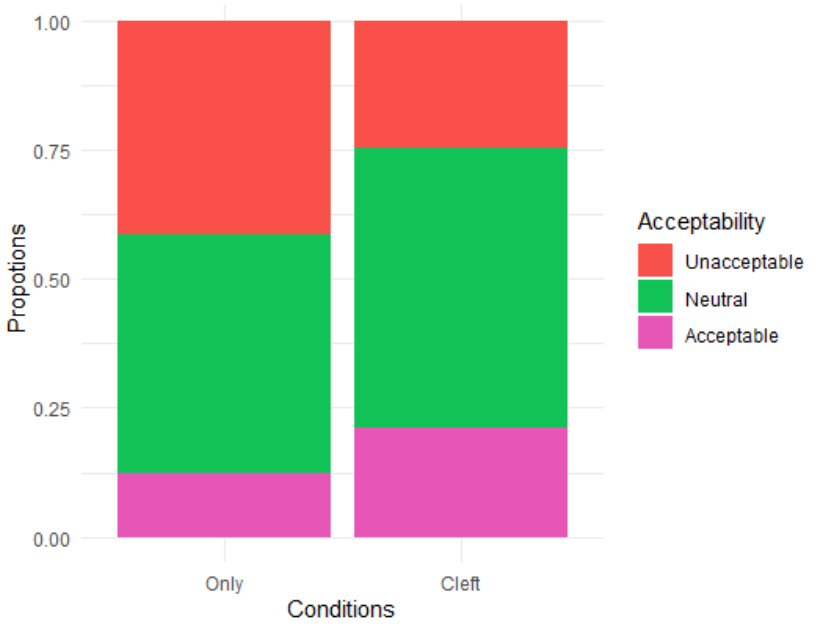


Figure 36 – Proportions of acceptability ratings across syntactic structure in AJT3.

We modelled another Pearson's chi-squared test to check the association between exhaustivity and acceptability, which we found to be significant as well (Figure 37). In fact, the exhaustivity seemed to be more influential for acceptability ($\chi^2(2) = 168.67$, $p < 0.00$) than the variation of syntactic structures.

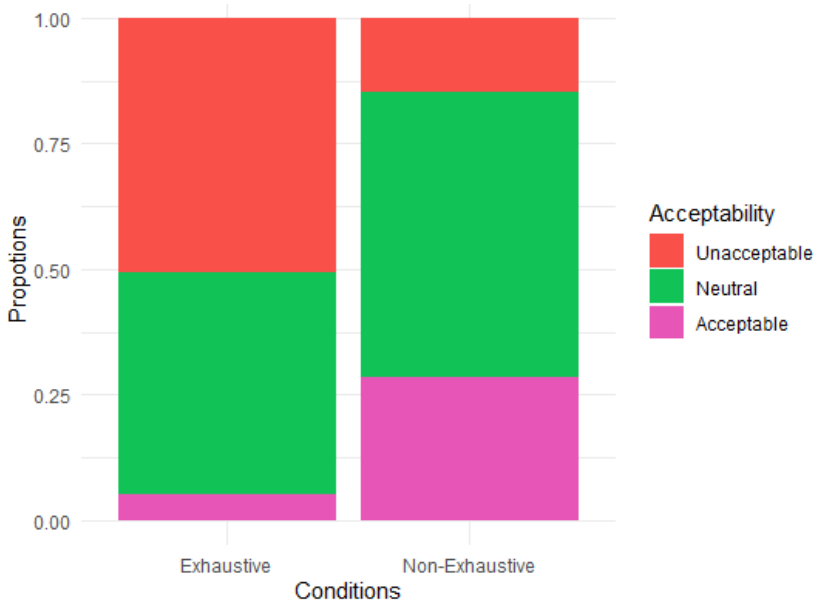


Figure 37 – Proportions of acceptability ratings across exhaustivity AJT3.

Finally, we proceeded with a CLMM (*clmm*) to examine the effects of syntactic structure and exhaustivity on acceptability ratings. This model (Table 14) estimates the cumulative log-odds of a higher response on the acceptability scale (DVs), in relation to the syntactic structure and the exhaustivity (IVs). The DV variable was treated as an ordered factor ranging from 1 (extremely unacceptable) to 5 (extremely acceptable), and the model included the interaction of both IVs (syntactic structure and exhaustivity), as fixed effects. We included intercepts for participants (n=40) and items (n = 48) as random effects.

		Answer	
Predictors	*Odds Ratios*	*CI*	*p*
1|2	0.00	0.00 – 0.01	**<0.001**
2|3	0.05	0.03 – 0.10	**<0.001**
3|4	0.18	0.10 – 0.31	**<0.001**
4|5	2.47	1.40 – 4.38	**0.002**
Structure [Only]	0.42	0.24 – 0.75	**0.003**
Exhaustivity [Non_exhaustive]	0.11	0.06 – 0.20	**<0.001**
Structure [Only] × Exhaustivity [Non_exhaustive]	0.64	0.29 – 1.40	0.264
Random Effects			
σ^2	3.29		
$\tau_{00\ Item}$	0.55		
$\tau_{00\ Participant}$	1.53		
ICC	0.39		
$N_{\ Participant}$	40		
$N_{\ Item}$	96		
Observations	897		
Marginal R^2 / Conditional R^2	0.246 / 0.539		

Table 14 – CLMM with interaction of the variables in AJT3.

The results revealed significant main effects of both syntactic structure (OR = 0.42, 95% CI [0.24, 0.75], p = 0.003) and exhaustivity (OR = 0.11, 95% CI [0.06, 0.20], p <0 .001). Compared to exhaustive clefts, 'only'-sentences decreased the acceptability ratings, and so did non-exhaustivity. The interaction between 'only' and non-exhaustivity had no effects (OR = 0.64, 95% CI [0.29, 1.40], p = 0.264), and this result was different from AJT1 ('...did too') but similar to AJT2 ('and also…').

4.1.3.6 Summary for all the AJTs

In summary, the observations for the AJTs 1, 2 and 3 were: 1) clefts were more likely to receive higher acceptability rates, especially in exhaustive sentences; 2) exhaustivity increased acceptability chances compared the violation of such semantic effect; 3) although 'only' and exhaustivity increase acceptability in every task when isolated, their interaction had a significant effect solely in the first task (AJT1). Importantly, regardless of the type of syntactic structure, all the AJTs showed a strong decrease in acceptability with non-exhaustivity. This result was expected, as the non-exhaustive conditions represented semantic violations. In BP, this effect was clearer than in English, with participants presenting an 89% of chances (OR = 0.11) for lower ratings of non-exhaustive sentences in AJT3.

As in AJT2 (OR = 0.85, 95% CI [0.54, 1.32], p = 0.466), no significant interaction between structure and exhaustivity was observed in AJT3 (OR = 0.64, 95% CI [0.29, 1.40], p = 0.264), indicating that the effect of exhaustivity was consistent across syntactic structures. Interestingly, AJT2 used the 'and also…' additive structure, that is syntactically and semantically similar to the '*e também*' structure in BP, used in AJT3.

Unlike the VP-ellipsis construction used in AJT1 ('...did too'), the additive structures used in AJT2 and AJT3 ('and also…'/'*e também*…') do not necessarily enforce a strict contrastive relation between alternatives. The VP-ellipsis requires the reconstruction of a parallel event and highlights a contrast between subjects (Carlson, 2001; 2015). On the other hand, the *gapping* additive structures ('and also…' are not necessarily perceived as contrastive from the beginning, which allows them a more flexible integration: this second proposition may be interpreted as a simple addition, rather than a direct alternative to the first. As a result, in the context of *gapping* sentences, an exhaustivity violation might affect acceptability globally, which would explain why it did not interact with syntactic structure.

Conversely, our results could suggest that VP-ellipsis makes the contrast between alternatives more explicit, which strengthens the dependency between 'only' and exhaustivity, yielding a significant interaction.

It is important to revisit the distinction between exhaustivity in 'only'-sentences and in clefts. Exhaustivity in 'only'-sentences has been argued to follow from their logical structure (König, 1992; Rooth, 1992; Kiss, 1998), while clefts do not necessarily encode exhaustivity in the same way, as their interpretation would depend mainly on semantic and pragmatic factors (Horn, 1981; Krifka, 2008). Building on this distinction, and considering that exhaustivity arises from contrastive focus (Rooth, 1992), we argue that the absence of interaction effects in AJT2 and AJT3 is not due to the nature of clefts *per se*, but probably to the type of relation established between propositions.

In the additive constructions used in these tasks ('and also…', '*e também…*'), the second clause can be integrated as an addition rather than as a strictly contrastive alternative. As a result, violations of exhaustivity affect acceptability uniformly, regardless of structure. In contrast, VP-ellipsis constructions (AJT1) promote a tighter contrast between alternatives, increasing the dependency between focus marking and exhaustivity, which establishes a significant interaction with 'only'.

4.2. Eye-tracking task in English

The acceptability tasks in English provided support for our hypothesis: exhaustivity violations in 'only'-sentences were judged to be less acceptable than the same violations in clefts. Furthermore, 'only'-sentences were consistently rated as less acceptable than it-clefts, regardless of exhaustivity, which indicates that the effect of exhaustivity was not uniform across the conditions, depending on the syntactic structure. In particular, violations of exhaustivity in 'only'-sentences led to a significant decrease in acceptability, while such disruptions were apparently more tolerated for clefts. Besides, there was a clear distinction between the results from the AJT1 ('... did too') and the ones for AJTs 2 and 3 ('and also…'/'e também…'). These results provide evidence to the proposition that the syntactic structure modulates how exhaustivity is interpreted and, therefore, semantic violations are perceived.

The stimuli for the AJTs were presented with no background context[27], and we exposed the methodological reasons for this in section 4.1. In sum, we wanted readers to rate trial sentences relying exclusively on intrinsic interpretive properties, rather than contextual information. For this purpose, presenting items with no previous context was not only enough, but actually more adequate.

Conversely, it is important to notice that the lack of context allows ambiguous interpretations, once focus-sensitive constructions are susceptible to context (Bard *et al*., 1996; Krifka, 2008). To address this limitation, we designed an eye-tracking experiment that presented context-sentences before each experimental stimulus. This created a different environment for the comprehension of it-clefts and 'only'-sentences, as readers would be forced to integrate the information of the target-sentences to what has been previously processed.

Along section 3, we have presented evidence to believe that focus adverbs such as 'only' associate to the truth-conditions of the sentence, whereas it-clefts convey their meaning because of pragmatic implicature (Rooth, 1992; Kiss, 1998; Krifka, 2008; Drenhaus *et al*., 2010; Carlson, 2013). Taking this discussion into account, we expected that context-sentences would have greater influence in clefts than in 'only'-sentences.

Therefore, we designed two eye-tracking tasks (one in English and one in BP) that would check our second general hypothesis: 2) exhaustivity violations in it-clefts are more sensitive to context than the same violations in 'only'-sentences. The details of these experiments are presented in the subsections below.

4.2.1. Materials

The eye-tracking reading task in English contained 120 experimental sets[28], plus 120 sets of fillers. There were also five (5) training-sets before the task started. All the sentences contained transitive verbs, and their subject-NPs were carefully chosen so as half of the sets

[27] It is worth mentioning that despite the lack of context-sentences previously to the experimental items of the AJTs, all of such items contained adverbial phrases in the beginning. These phrases would modulate the information about the event and the people involved in it, to a certain extent. Therefore, participants would be forced to accommodate sentences with a complicated focus structure using exclusively the idea provided by the adverbial phrases within them, with no surrounding discourse context (Carlson, 2014).

[28] See Appendix D for all the items.

(60) presented female subject-NPs, and half of them presented male ones, to minimise gender biases.

Each set contained two sentences: 1) a context-sentence, which necessarily mentioned two subject-NPs in a cooperative context; 2) a target-sentence, that appeared in six different conditions. The target-sentences, that is, the experimental items' conditions were: a) baseline with two object-NPs, conveying exhaustivity (Baseline [exh+]) for the subject-NP; b) baseline with two subject-NPs, non-exhaustive (Baseline [exh-]); c) 'only'-sentences, conveying exhaustivity (Only [exh+]); d) 'only'-sentences, but non-exhaustive (Only [exh-]); e) it-cleft, conveying exhaustivity (Cleft [exh+]); and f) it-cleft, but non-exhaustive (Cleft [exh-]). There is an example of it (13):

(13) Context-sentence: Anna and Olivia run a busy bakery in London.

a. Yesterday **Anna** brewed coffee and tea too. (baseline; [+exh])
b. Yesterday **Anna** brewed coffee and **Olivia** did too. (baseline; [-exh])
c. Yesterday **only Anna** brewed coffee and tea too. ('only'; [+exh])
d. Yesterday **only Anna** brewed coffee and **Olivia** did too. ('only'; [-exh])
e. Yesterday **it was Anna who** brewed coffee and tea too. (it-cleft; [+exh])
f. Yesterday **it was Anna who** brewed coffee and **Olivia** did too. (it-cleft; [-exh])

We believe the reasons why items 13(a, c, e) are so-called 'exhaustive' and items 13(b, e, f) are 'non-exhaustive' have already been overtly discussed so far. However, there are still a few important considerations about our items.

Firstly, it is clear that this experiment used the '…did too' structure, not the 'and also…' one. That was because of what we have exposed about the first one being associated with a VP-ellipsis and the latter with a *gapping* phenomenon (Clifton; Bock; Rado, 2001; Johnson, 2009; Da Silva; Fonseca, 2021). If we consider that *gapping* structures might be more complex to process than VP-ellipses (Johnson, 2009), any results indicating difficulties towards the items could be related to the very nature of the *gapping*, which would be difficult to disentangle from difficulties emerged by exhaustivity violations. This could possibly mask our results, so '…did too' was probably a better choice over 'and also…' to avoid any kind of biases that the latter could represent[29]. Additionally, the results of AJT1 (VP-ellipsis; '...did

[29] We did not have the same concerns about the 'and also...' structure with the AJT2, because task 2 was conducted *offline*. Eye-tracking experiments are online, which means that they are expected to find difficulties of

too') showed an effect of interaction between exhaustivity and syntactic structure, which provides evidence that exhaustivity is processed in different ways for these types of structures, making the VP-ellipsis more interesting to investigate, as it had already presented some results[30].

Secondly, we produced all the items in collaboration with native-speakers of English of the UK, at the location where we applied the experiment. Thus, we guaranteed orthographical, semantical and syntactic correction to all items in this task.

Lastly – and most importantly – we have included two baseline conditions, in comparison to the acceptability tasks. We did that to create opportunities to look for effects intrinsic to 'only' and/or clefts isolated, without having to compare one another. Without the neutral structures (baselines), with no 'only' or clefts, we necessarily compared 'only'-sentences with it-clefts, and so we were unable to analyse the effects of the focus adverb and/or the clefts alone.

All sentences (context-sentences and target-sentences) were presented in full on the screens. Nevertheless, each target-sentence was segmented into areas of interest (AOI) on SR Research Experiment Builder (SR Research, n.d.), which gave us a table of results in which parameters' data was given per AOI.

The segmentation (Table 15) that we used was based on syntactic structure and semantic role, and it was very important for our data analysis.

Adverb (AOI 1)	**Focus marker (AOI 2)**	**Subject (AOI 3)**	**Who (AOI4)**	**Verb (AOI 5)**	**Object (AOI 6)**	**Conjunction (AOI 7)**	**Critical segment (AOI 8)**
Yesterday		**Anna**		brewed	**coffee**	and	**tea too.**
Yesterday		**Anna**		brewed	**coffee**	and	**Olivia did too.**
Yesterday	only	**Anna**		brewed	**coffee**	and	**tea too.**
Yesterday	only	**Anna**		brewed	**coffee**	and	**Olivia did too.**
Yesterday	it was	**Anna**	who	brewed	**coffee**	and	**tea too.**
Yesterday	it was	**Anna**	who	brewed	**coffee**	and	**Olivia did too.**

processing for which comprehenders are unconscious about too (Ferreira; Clifton, 1986; Traxler, 2013). By minimising difficulties that were unrelated to what we were testing, we imagined that the results would become clearer for what really mattered.

[30] On the other hand, results showed no such effect of interaction between these variables in AJT2, making the additive *gapping* structure ('and also…') a less obvious alternative to investigate syntactic effects on exhaustivity.

Table 15 – Areas of interest (AOIs) of the eye-tracking experiment.

4.2.2. Method and Apparatus

This study was developed at the School of Psychology of the University of Leicester – UoL (UK). The research was approved by the *Health, Biological and Psychological Sciences Research Ethics Committee* in the School of Psychology and Vision Sciences at the University of Leicester (ID: 2859).

We conducted the experiment using an EyeLink 1000 Plus system (SR Research Inc., Ontario, Canada), at a sampling rate of 1000 Hz. The experiment was designed on Experiment Builder. As there were 6 conditions, there were 6 versions of the experiment (counterbalances 1, 2, 3, 4, 5 and 6). Every time a participant arrived, we manually selected the version of the experiment that we would like them to take. This was the procedure until we had the same number of participants for each version. All stimuli were randomised on Experiment Builder, and counterbalanced across the 6 lists, following a Latin square design (Abbuhl; Gass; Mackey, 2013).

Once participants came to the laboratory, we first asked them to read and sign a Consent Form (Appendix F) containing general details of the experiment and ensuring that personal data and data produced by them would remain anonymous. Then, we gave them a brief explanation about the task and about the buttons that they should press when they had to. Participants were told that they should read each sentence naturally for comprehension. We started testing the set-up calibration before and after each sentence, and recalibrated the equipment whenever it was necessary throughout the session.

A binocular mount and a chin-and-forehead rest minimised head movement and ensured accurate gaze tracking. Participants sat approximately 65 cm from a BenQ high-resolution monitor (1920×1080 resolution, 144-Hz refresh rate), where sentences appeared in black ‘Courier New’ font (14-point) on a light-gray background (Rayner; Pollatsek, 1989).

Each trial began with short context-sentences (one line, as in example 5) displayed on the screen. After having read the context-sentences, participants should press a white button, in the centre of a SR Research Button Box (SR Research Inc., Ontario, Canada). That would show participants the next screen, with the experimental items (target-sentences). Participants had a practice section (with 5 sets of sentences), and then proceeded to the trial.

For 40% of the stimuli, that is, 96 sets of sentences (48/120 experimental items and 48/120 fillers), we included a yes/no comprehension question, such as: 'Did Anna brew milk?'[31]. Participants should press a red button (on the left) for *No* responses, or a green button (on the right) for *Yes* responses, and a white button in the center used to proceed to the next sentence. We made this to measure (and retain) participants' attention during the trial. As there were many stimuli (240 experimental ones), we divided the trial into three sessions with 80 stimuli each. During the pauses, participants saw a screen with the instruction to press the white button (in the centre) whenever they were ready to continue.

4.2.3. Independent Variables

The IVs of this task were: 1) (non-) exhaustivity; 2) syntactic structure; and 3) area of interest (AOI). It is worth noting that in this task we included two baseline conditions, which gave us two conditions of exhaustivity (exhaustive/non-exhaustive) and three conditions of syntactic structure (baseline/'only'-sentences/it-clefts) and eight areas of interest (AOI). Considering that not all eight were informative for our purposes, we analysed data of four different AOIs, as we will discuss in the Results section.

4.2.4. Dependent Variables

The EyeLink 1000 Plus system registers many different parameters, which is why our biggest challenge was to determine which measures would be more interesting (and more informative) for our research. We expected some complexity during the processing of non-exhaustive sentences, and we were particularly interested in comparing these parameters across the different areas of interest (AOIs). With regard to that, we favoured parameters that typically show difficulty of processing (Rayner, 1998; Clifton; Staub; Rayner, 2007; 2009; Fonseca; Maia, 2022). Accordingly, our analysed DVs were: 1) regression-in count; 2) regression-out count; 3) regression-path duration; 4) fixation count; and 5) first-fixation duration.

Regressions-in count (IA_REGRESSION_IN_COUNT, on EyeLink 1000) corresponds to the number of regressions into a given AOI – that is, how often the reader moved their gaze back to that region, after already having left it. Regressions are eye

[31] This was not extracted from the stimuli of the experiment, as the set of sentences (5) did not have any comprehension questions to refer to.

movements against the natural reading direction, so they usually indicate difficulties of processing or integration problems (Rayner, 1998; Clifton; Staub; Rayner, 2007). Specifically, a high number of regressions into an AOI suggests revaluation or reanalysis of the information presented in that region.

On the other hand, regressions-out count (IA_REGRESSION_OUT_COUNT, on EyeLink 1000) refer to eye movements in which the reader moves their gaze away from the current AOI to a previous region of the text, so they indicate that the reader is leaving the current region to look back to previous regions. This might be due to processing difficulty, lack of integration with previous context, or the need to re-evaluate earlier material (Rayner, 1998; Clifton; Staub; Rayner, 2007). Because they reflect integration processes and reanalysis, regressions-out are considered late-stage measures in the time course of reading, in contrast to regressions-in. High regressions-out rate typically associated with difficulties that readers could not resolve in the AOI that they are currently reading, meaning a need to retrieve information from earlier segments of the sentence.

Regression-path duration captures the total duration of all fixations, from the moment that the reader enters a given area for the first time until they move past it to the right (next area), including any regressions to earlier parts of the sentence. This metric reflects proportional difficulties of processing that causes the reader to spend time reanalysing earlier segments before continuing to read.

Fixation count is the total number of fixations made within a given AOI during the reading of a sentence or text segment. It captures the overall amount of visual attention devoted to that area and is commonly used as a measure of processing effort or difficulty (Rayner, 1998). Higher fixation counts usually imply that the reader needed to reexamine or spend more time on the region, indicating greater cognitive demand.

First-fixation duration refers to the time of the very first fixation that a reader makes on a specific area of interest (AOI) during reading. It reflects early visual and cognitive processing of that region, often indexing initial word recognition and lexical access (Rayner, 1998; Clifton, Staub; Rayner, 2007). Shorter first fixation durations typically indicate easier processing, while longer durations suggest increased processing difficulty or complexity.

All these metrics were analysed and reported in the Results section (4.2.6).

4.2.5. Participants

Forty (40) undergraduate students of the School of Psychology of the University of Leicester participated in this eye-tracking experiment. We analysed data from 36 people, after counterbalancing the results table. All participants were native speakers of English, with ages between 18 and 21 years, and took part in the experiment in exchange of obligatory credits for their academic year at the University.

Participants signed up for the experiment online, on the University's system, by selecting a date and time slot to schedule their visit to the laboratory. On the system we asked people who had any kind of reading impairments, such as dyslexia or uncorrected/uncorrectable vision deficits, not to join the experiment.

4.2.6. Results

The results of this experiment were analysed on *RStudio* (R Core Team, 2025).

We had previously prepared the results' table, by removing data from participants that were not able to answer at least 80% of the yes/no comprehension questions correctly – that is, 39/48 questions. We made this to measure (and retain) participants' attention during the trial. Additionally, we assured that the table was counterbalanced, so we had the same number of participants (n = 36) for each version of the experiment (6 participants per version). The balanced results table contained 24,185 observations.

In the present analysis, different types of eye-tracking measures required distinct statistical approaches, due to the nature of their distributions. We analysed measures expressed as discrete data (regression-in counts, regression-out counts, and fixation counts) using generalised linear mixed-effects models (GLMMs) with a Poisson distribution (Bates *et al*., 2015) – *glmer* function from the *lme4* package, on *RStudio*. This choice reflects the non-normal and discrete nature of these variables, which often violate the assumptions of linear models (Vasishth *et al*., 2018).

Contrastively, for continuous duration-based measures (such as regression-path duration), we applied log transformations to improve the distributional characteristics of the data, and then used linear mixed-effects models (LMMs) – using *lmer* function on RStudio. This approach aligns with current practices in Psycholinguistic research using eye-tracking, as

time data is positively skewed[32] in eye-tracking experiments, and rarely follow a normal distribution (Baayen; Milin, 2010; Staub, 2015). Moreover, recent work has shown that untransformed LMMs may yield inflated effect sizes, while GLMMs or log-transformed LMMs provide more conservative and robust estimates (Godwin; Lee; Drieghe, 2025). Therefore, the combination of GLMMs for count data and log-transformed LMMs for duration measures tends to increase the accuracy of the inferences.

All of the EyeLink measures on the results' table were given by areas of interest (AOI). Let us recall which were our AOIs (Table 17) in this experiment: 1) adverb; 2) focus marker; 3) subject-NP; 4) 'who' (it-clefts); 5) verb; 6) object-NP; 7) 'and'; 8) critical/final segment. Not all sentences contained lexical material in all AOIs (*e.g.*: baselines did not have AOI 2; 'only'-sentences and baselines did not have AOI 4, so our analyses did not follow the same pattern for all AOIs or all parameters. In the subsections below we will describe our analysis for AOI. We modeled GLMMs for regression-in, regression-out counts and fixation counts and LMMs for regression-path duration and first-fixation duration. Participants and items were included as random effects in every model. There were four different levels for conditions in AOI2, namely: exhaustive cleft (Cleft [exh+]) – that was the intercept – non-exhaustive clefts (Cleft [exh-]) exhaustive 'only'-sentences (Only [exh+]) and non-exhaustive 'only'-sentences (Only [exh-]). These levels represented the different combinations for structure (it-cleft/'only') and exhaustivity (exhaustive/non-exhaustive).

For each AOI, we tested the main effects of the independent variables (structure and exhaustivity) first, then proceeded with further analysis, depending on our findings. The results and analyses presented in this section are organised by AOI.

4.2.6.1. Area of the focus markers (AOI2)

The AOI2 contained the focus marker – 'only' or the cleft ('it was') – which associates to the subject-NP and introduces presuppositions regarding exhaustivity. Because of that, it was expected that this area (AOI2) received regressions, once participants encountered the subsequent challenges for the semantic predictions.

---

[32] Skeweness is a characteristic of asymmetrically distributed data, which is very common for reading measures of an experiment, as they can vary a lot, according to participants, items and/or conditions. A positive skewness (right-skewness) means that the histogram that represents this data has a long right tail, indicating that the outliers concentrate on the right side of the mean. This is the same as saying that the outliers are extreme values that are higher than the mean value. Reading measures tend to be positively skewed, as it is more common for outliers to be higher than the mean, and not lower than the mean – the later constitutes a negative skewness (or left-skewness), and it is the opposite of the former (Negi; Mitra, 2020).

A descriptive analysis showed that exhaustive conditions elicited more regressions into AOI2 than non-exhaustive conditions (Figure 38):

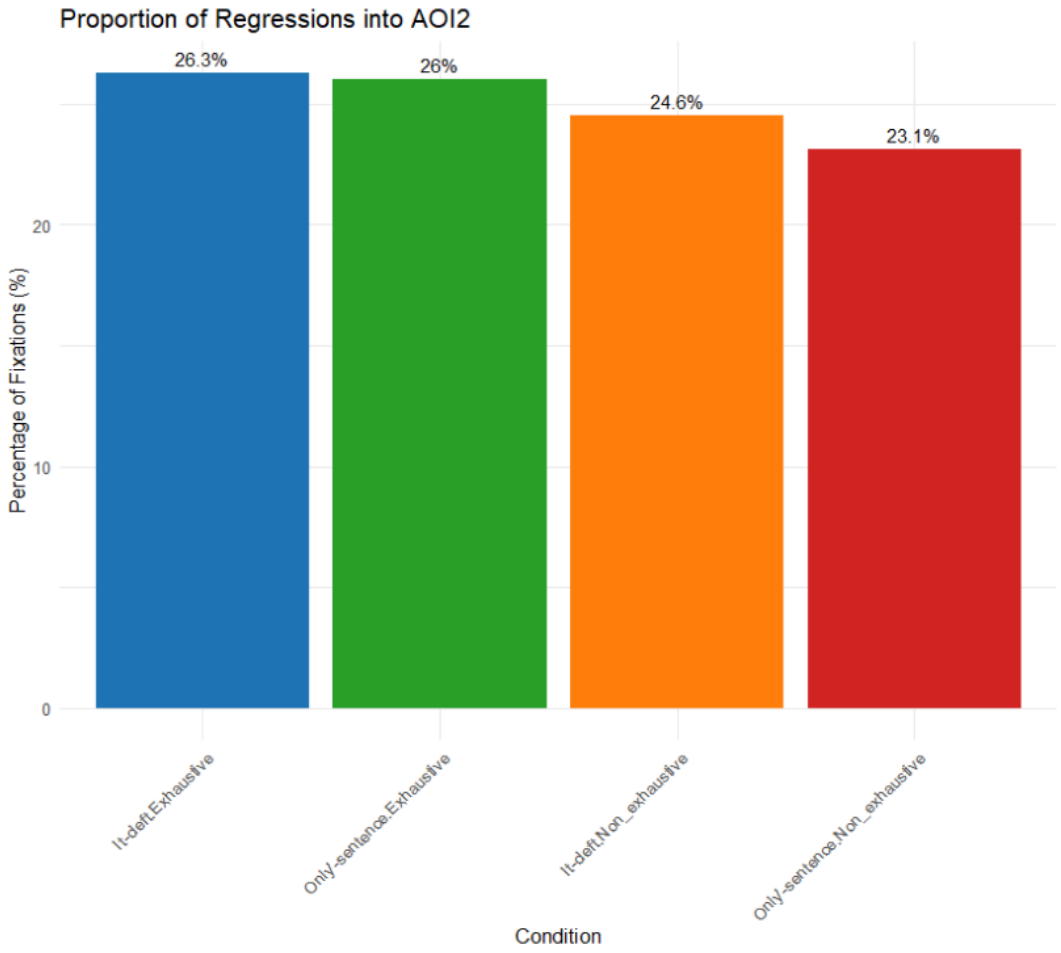


Figure 38 – Proportions of regressions into AOI2 per condition.

We started our inferential analysis checking for main effects, considering the independent variables as fixed effects in the GLMM. Results show an effect of structure (isolated) and no effect of exhaustivity (Table 16). The results concerning the structure as a variable was not surprising, as AOI2 contained the focus markers, so these segments of the sentences were essentially different in syntactic structure. However, we expected to find more regressions into this area for non-exhaustive sentences, as they presented semantic violation and therefore potentially more complex processing.

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.54	0.45 – 0.63	**<0.001**
Structure [Only'-sentence]	1.19	1.02 – 1.39	**0.024**
Exhaustivity [Non_exhaustive]	0.95	0.82 – 1.11	0.552
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	1.00	0.80 – 1.25	0.996
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.14		
$N_{\ Participants}$	36		
$N_{\ Item}$	479		
Observations	2125		

Table 16 – GLMM for regression-in count into AOI2.

Regarding regression-path duration means in AOI2 – an initial descriptive analysis – we found that regressions into clefts (‘it was’) seemed to be longer than regressions into the adverb ‘only’ (Figure 39). It is important to consider, though, that this segment (AOI2) was simply longer in clefts (six characters) than in ‘only’-sentences (four characters), so this difference of length could explain longer regressions into clefts.

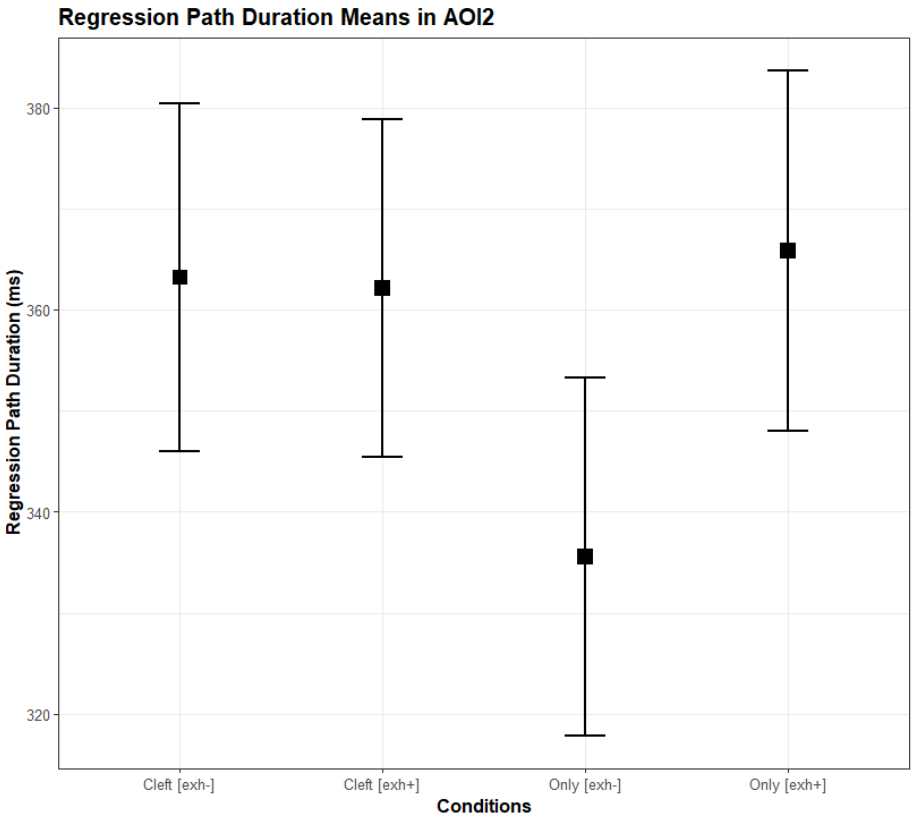


Figure 39 – Regression-path duration means into AOI2.

However, we modeled a LMM (Table 17) that revealed a significant effect for the interaction of variables (‘only’-sentences x non-exhaustivity), and not for the structure isolated, which means that the length of the area, alone, did not influence regression-path duration means.

Figure 40 illustrates that the effect of (non-)exhaustivity depends on the effect of structure.

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.73	5.66 – 5.80	**<0.001**
Structure [Only'-sentence]	0.02	-0.04 – 0.08	0.559
Exhaustivity [Non_exhaustive]	0.01	-0.05 – 0.07	0.744
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	-0.10	-0.19 – -0.02	**0.021**
Random Effects			
σ^2	0.25		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.04		
ICC	0.13		
$N_{\ Participants}$	36		
$N_{\ Item}$	479		
Observations	2084		
Marginal R^2 / Conditional R^2	0.004 / 0.131		

Table 17 – LMM for regression-path duration in AOI2.

There was a higher amount of fixations in AOI2 of 'only'-sentences than in AOI2 of clefts (Figure 41). However, after analysing this metric with a GLMM (Table 19), we found no significant effects. Likewise, we found no effects in first-fixation durations (Table 20).

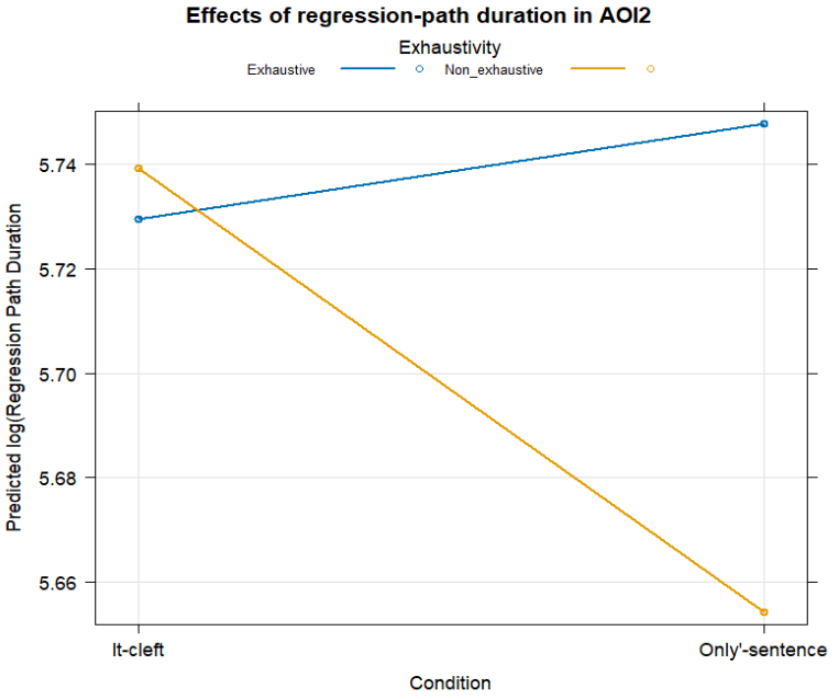


Figure 40 – Regression-path duration means into AOI2.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	1.47	1.36 – 1.59	**<0.001**
Structure [Only'-sentence]	0.96	0.87 – 1.06	0.427
Exhaustivity [Non_exhaustive]	0.97	0.88 – 1.06	0.487
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.97	0.84 – 1.12	0.706
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
$N_{Participants}$	36		
N_{Item}	479		
Observations	2110		

Table 18 – GLMM for fixation counts in AOI2.

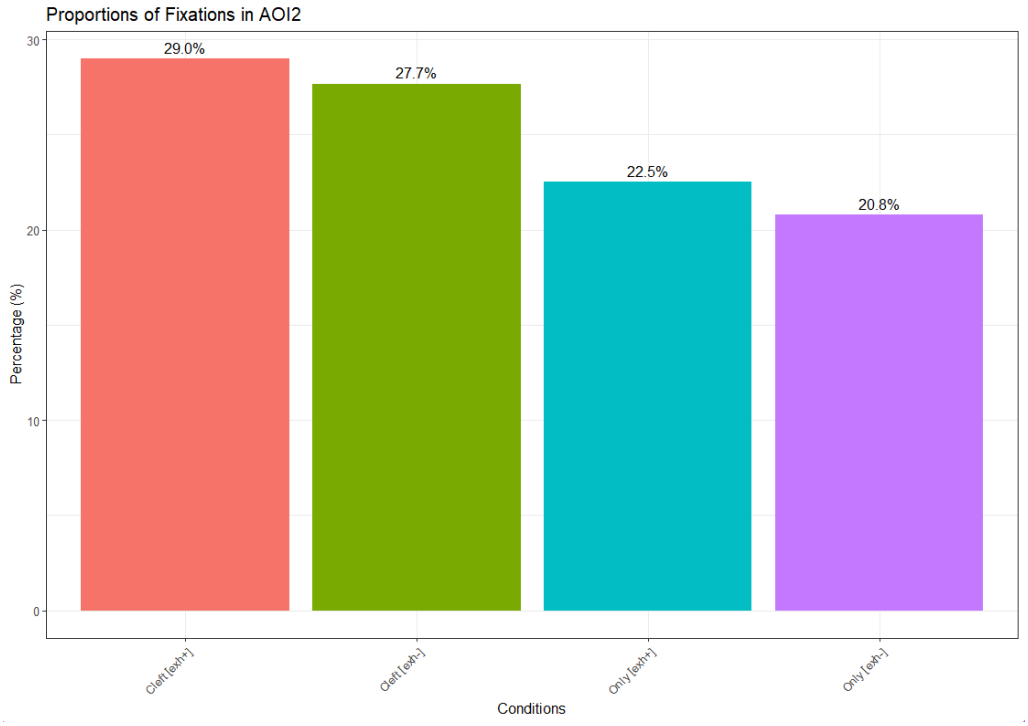


Figure 41 – Proportions of fixation count in AOI2.

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.20	5.15 – 5.24	**<0.001**
Structure [Only'-sentence]	0.03	-0.01 – 0.07	0.121
Exhaustivity [Non_exhaustive]	0.01	-0.02 – 0.05	0.542
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.01	-0.04 – 0.07	0.634
Random Effects			
σ^2	0.09		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.13		
$N_{\ Participants}$	36		
$N_{\ Item}$	479		
Observations	2034		
Marginal R^2 / Conditional R^2	0.004 / 0.135		

Table 19 – LMM for first-fixation duration in AOI2.

4.2.6.2. Area of the subject-NP (AOI3)

This area contains the subject-NP (*e.g.*, Anna), which was also the scope of the clefts/'only'. As expected, AOI3 was particularly sensitive to regressions-in, because this was the area that contained a focused piece of information (subject-NP).

We observed more regressions into AOI3 for the baseline conditions (Figure 42), which might have been exactly because these were the conditions that did not contain any focus markers (clefts or adverbs) – consequently, no AOI2. Because of that, people who needed to revisit subject-like information would regress directly to AOI3. Accordingly, we can say that this result was not surprising: regressions into the subjects were divided between AOIs 2 and 3 for conditions that contained focus markers ('it was / Anna' 'only / Anna'), whilst in baselines they were concentrated in AOI3.

Additionally, the descriptive analyses for regressions into AOI3 (Figure 42) shows that 'only'-sentences received a higher percentage of regressions-in than clefts:

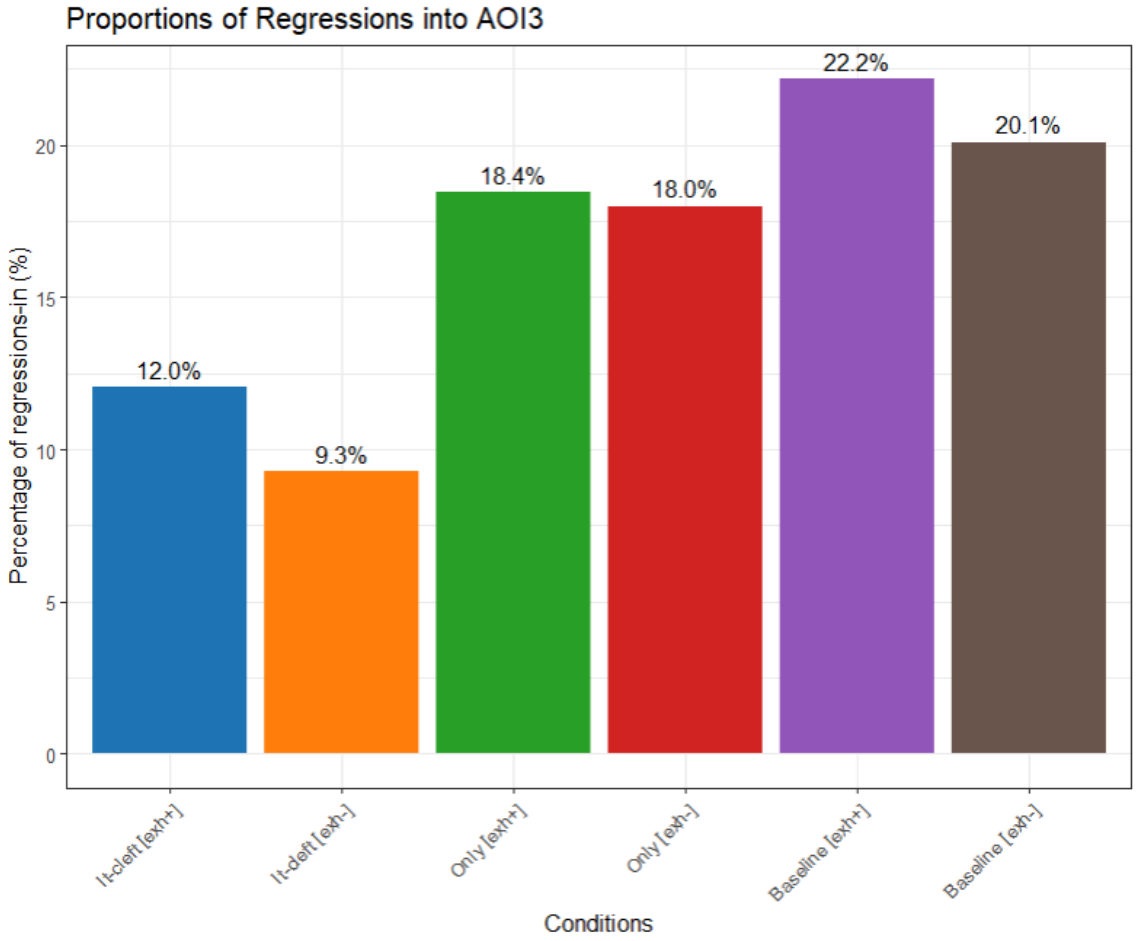


Figure 42 – Proportions of regressions into AOI3 per condition.

The GLMM (Table 21) showed significant effects for structure and exhaustivity, isolated and/or combined. This means that both structures ('only' and baselines) increased the likelihood of regressions into AOI3 (IRR = 1.35, 95% CI [1.13, 1.62], p = 0.001; IRR = 1.67, 95% CI [1.41, 1.99], p <0.001), regardless of exhaustivity conditions, when compared to clefts. Moreover, 'only' combined with non-exhaustivity also had a (marginally) significant effect to increase regressions-in (IRR = 1.31, 95% CI [1.01, 1.71], p = 0.043). Interestingly - and unexpectedly - non-exhaustivity (isolated) decreased the likelihood for regressions into AOI3 (IRR = 0.75, 95% CI [0.27, 0.40], p = 0.009) for clefts.

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.33	0.27 – 0.40	**<0.001**
Exhaustivity [Non_exhaustive]	0.75	0.61 – 0.93	**0.009**
Structure [baseline]	1.67	1.41 – 1.99	**<0.001**
Structure [Only'-sentence]	1.35	1.13 – 1.62	**0.001**
Exhaustivity [Non_exhaustive] × Structure [baseline]	1.28	0.99 – 1.66	0.063
Exhaustivity [Non_exhaustive] × Structure [Only'-sentence]	1.31	1.01 – 1.71	**0.043**
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Subject}$	0.19		
$N_{Subject}$	36		
N_{Item}	719		
Observations	3472		

Table 20 – GLMM for regression-in count comparing conditions into AOI3.

A descriptive analysis showed the highest means of regression-path duration for baselines and the lowest means for clefts (Figure 42). Later, a LMM (Table 20) demonstrated that this effect of structure for the baselines was extremely significant (β = 0.15, 95% CI [0.08, 0.21], p < 0.001).

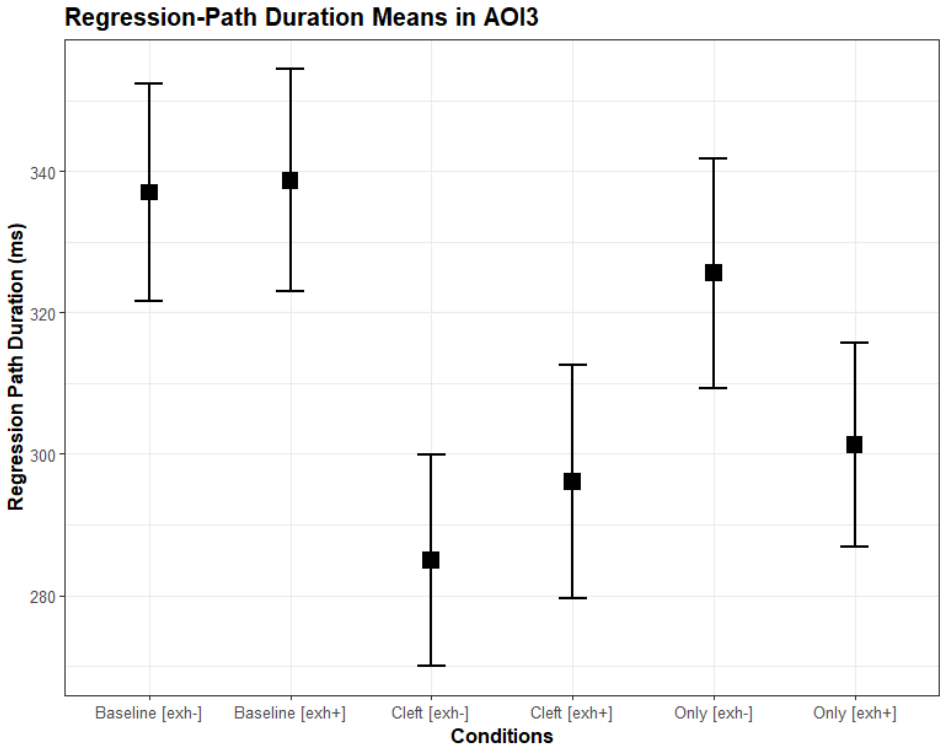


Figure 43 – Regression-path duration means into AOI3.

The same LMM (Table 21) also presented an interactive positive effect of non-exhaustivity and 'only', representing an increase of the likelihood in regressions-in there were slightly longer than the former (β = 0.09, 95% CI [0.00, 0.18], p = 0.040). In other words, the presence of 'only' and the non-exhaustivity potentially contributed to longer regressions into AOI3. These results match what we have observed for the focus marker area (AOI2), representing a significant effect of interaction between the two variables.

In conclusion, baselines and 'only'-sentences were likely to receive more and longer regressions into the subject-NP area (AOI3) than clefts. Additionally, non-exhaustivity influenced a potential decrease in regressions-in and regression-path duration, which was contrary to what we expected, since violations of exhaustivity are interpreted as difficulties during comprehension.

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.52	5.47 – 5.58	**<0.001**
Structure [baseline]	0.15	0.08 – 0.21	**<0.001**
Structure [Only'-sentence]	0.04	-0.02 – 0.11	0.167
Exhaustivity [Non_exhaustive]	-0.03	-0.09 – 0.04	0.417
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.04	-0.05 – 0.13	0.406
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.09	0.00 – 0.18	**0.040**
Random Effects			
σ^2	0.27		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.04		
$N_{\ Participants}$	36		
$N_{\ Item}$	719		
Observations	3294		
Marginal R^2 / Conditional R^2	0.017 / 0.058		

Table 21 – LMM for regression-path duration in AOI3.

An analysis of proportions for fixation counts (Figure 43) and a GLMM (Table 22) showed that the pattern observed for regressions into the subject-NP area (AOI3) for 'only'-sentences repeated on fixation proportions, in a way that 'only'-sentences received significantly more fixations (Figure 43; Table 22) than exhaustive clefts.

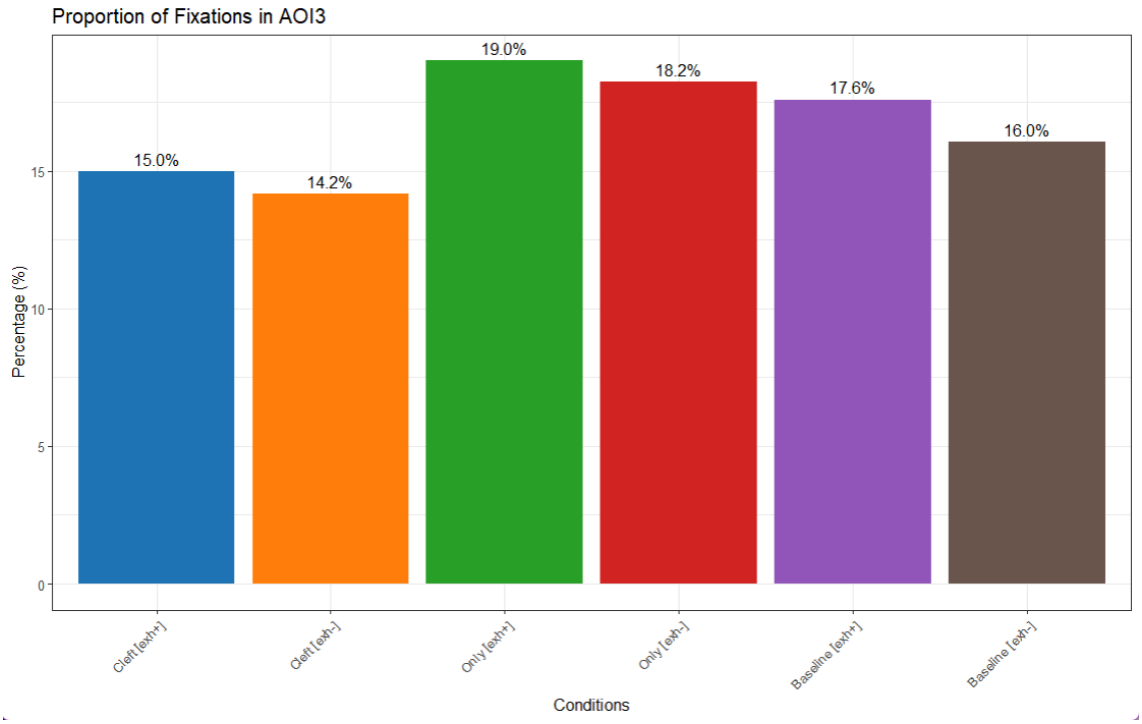


Figure 44 – Proportions of fixations in AOI3.

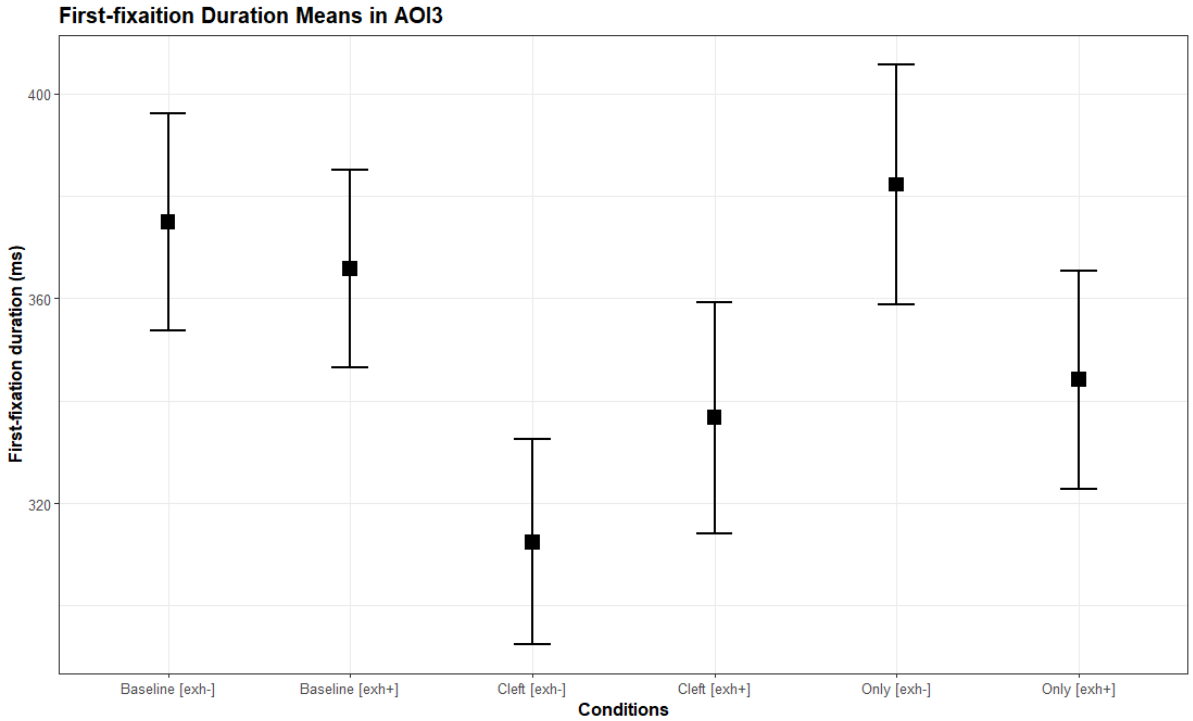


Figure 45 – First-fixation duration means in AOI3.

Moreover, our findings suggest that the adverb 'only' modulates processing at a very early stage, as first-fixation duration means were significantly higher for these structures, in comparison to clefts (Figure 44; Table 23).

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	1.41	1.30 – 1.52	**<0.001**
Structure [baseline]	1.05	0.96 – 1.16	0.302
Structure [Only'-sentence]	1.12	1.02 – 1.24	**0.016**
Exhaustivity [Non_exhaustive]	0.92	0.83 – 1.02	0.130
Structure [baseline] × Exhaustivity [Non_exhaustive]	1.04	0.90 – 1.19	0.592
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	1.06	0.93 – 1.21	0.400
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
$N_{\ Participants}$	36		
$N_{\ Item}$	719		
Observations	3448		

Table 22 – GLMM for fixation count in AOI3.

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.23	5.19 – 5.27	**<0.001**
Structure [baseline]	-0.00	-0.04 – 0.03	0.797
Structure [Only'-sentence]	0.04	0.00 – 0.08	**0.030**
Exhaustivity [Non_exhaustive]	-0.01	-0.04 – 0.03	0.788
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.03	-0.02 – 0.08	0.229
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.03	-0.03 – 0.08	0.334
Random Effects			
σ^2	0.09		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.09		
$N_{\ Participants}$	36		
$N_{\ Item}$	719		
Observations	3437		
Marginal R^2 / Conditional R^2	0.006 / 0.096		

Table 23 – LMM for first-fixation duration in AOI3.

4.2.6.3. Area of the object-NP (AOI6)

As one advances their reading of a sentence, regression-in counts generally diminishes, reflecting a reduction in chances of reinspection into later segments. In spite of

that, AOI6 comprises the object noun phrase (*e.g.*, 'coffee' in Table 15) that precedes the critical segment, and so it could become the target of reanalysis when readers encountered an unexpected continuation (semantic violation) in non-exhaustive sentences.

Descriptive analyses show that the cleft-conditions had the lowest proportions of regressions into AOI6, while 'only'-sentences and baselines were similar in terms of regression-in counts (Figure 45):

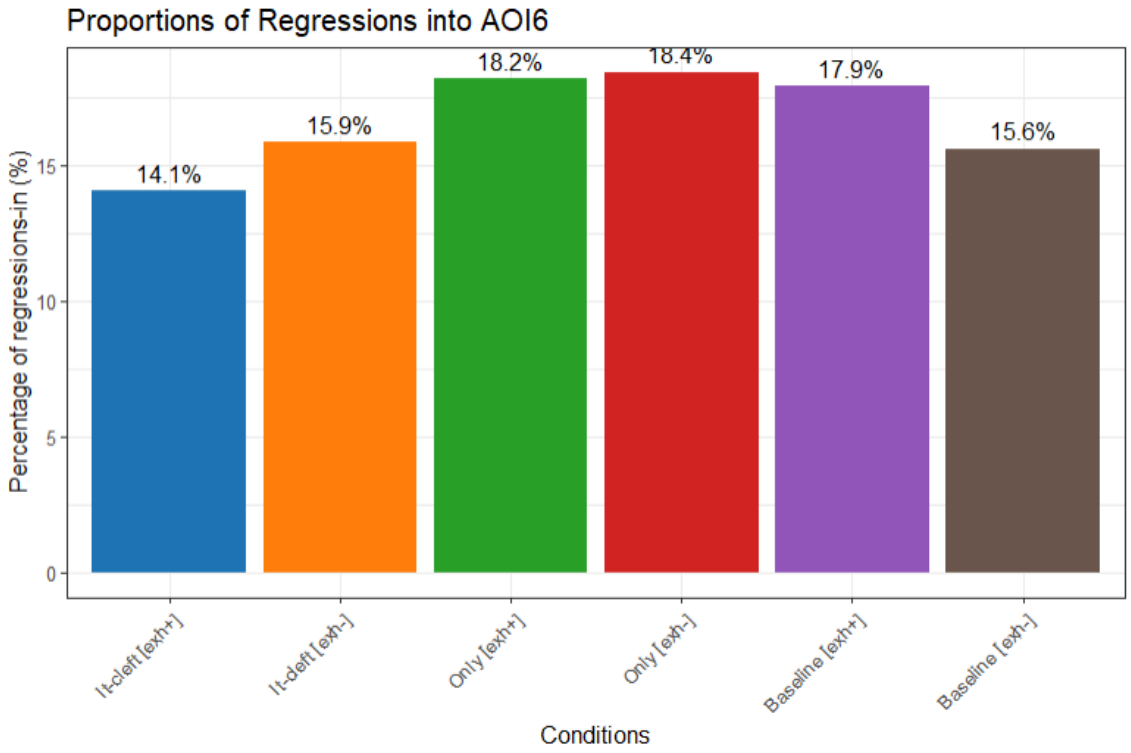


Figure 46 – Proportions of regressions into AOI6 per condition.

Regression-in analysis (GLMM – Table 24) revealed a main effect of structure, isolated – baselines (IRR = 1.53, 95% CI [1.28, 1.84], $p < 0.001$) and 'only'-sentences (IRR = 1.52, 95% CI [1.26, 1.82], $p < 0.001$) – and no effect of interaction between structure and exhaustivity (IRR = 0.82, 95% CI [0.63, 1.05], $p = 0.116$). In conclusion, 'only'-sentences and baselines, without considering exhaustivity conditions, increased the chance of regressions into the object-NP area, when compared to exhaustive clefts (intercept).

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.24	0.19 – 0.31	**<0.001**
Structure [baseline]	1.53	1.28 – 1.84	**<0.001**
Structure [Only'-sentence]	1.52	1.26 – 1.82	**<0.001**
Exhaustivity [Non_exhaustive]	1.18	0.97 – 1.42	0.100
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.90	0.70 – 1.15	0.386
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.82	0.63 – 1.05	0.116
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Subject}$	0.34		
$N_{Subject}$	36		
N_{Item}	720		
Observations	4039		

Table 24 – GLMM for regression-in count in AOI6.

Conversely, the regression-path duration descriptive analysis revealed a different pattern, in which the regression-path duration means for clefts were higher than for 'only'-sentences or for baselines (Figure 46). Nonetheless, the tested model showed no main effects of structure or exhaustivity for this metric (Table 23), so the differences of regression-path duration (Figure 46) were not significant, hence not interpretive.

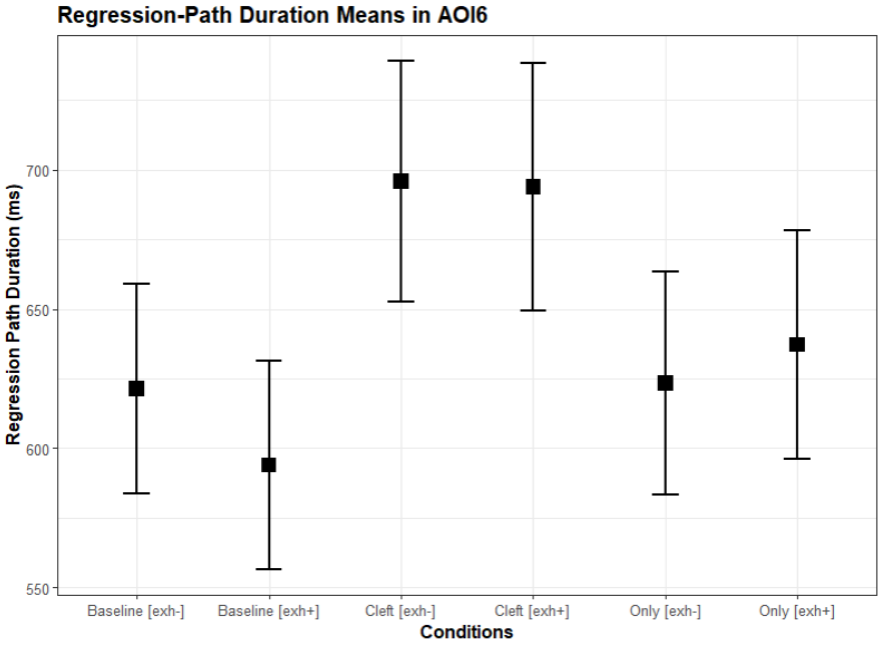


Figure 47 – Regression-path duration means in AOI6.

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.20	6.10 – 6.30	**<0.001**
Structure [baseline]	-0.11	-0.23 – 0.01	0.074
Structure [Only'-sentence]	-0.05	-0.17 – 0.07	0.441
Exhaustivity [Non_exhaustive]	0.04	-0.08 – 0.16	0.553
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.01	-0.16 – 0.18	0.919
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	-0.06	-0.24 – 0.11	0.458
Random Effects			
σ^2	0.51		
$\tau_{00\ Item}$	0.14		
$\tau_{00\ Participants}$	0.03		
ICC	0.25		
$N_{Participants}$	36		
N_{Item}	720		
Observations	4092		
Marginal R^2 / Conditional R^2	0.003 / 0.248		

Table 25 – LMM for regression-path duration in AOI6.

Figure 47 shows very homogenous proportions of fixation counts across conditions, and a GLMM (Table 26) confirmed that there were no early effects of exhaustivity or structure in the object-NP area (AOI6), as there were no significant differences in fixation count. Likewise, we found no effects for first-fixation duration (Table 27), and the exclusive effects for regressions into AOI6 suggest that this area was more important to later processing - probably confirmatory checks, mainly for 'only'-sentences (and baselines).

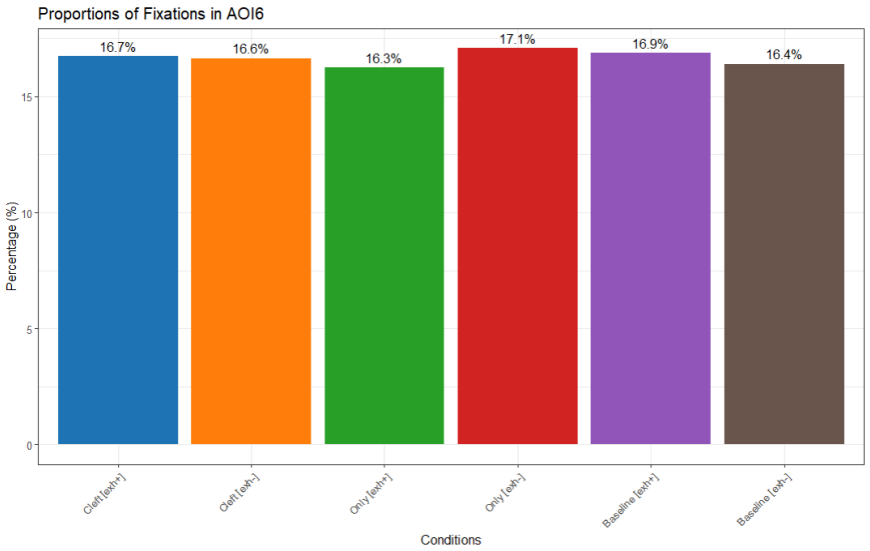


Figure 48 – Proportions of fixations count in AOI6 per condition.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	2.00	1.88 – 2.12	**<0.001**
Structure [baseline]	1.00	0.92 – 1.08	0.951
Structure [Only'-sentence]	1.00	0.92 – 1.08	0.928
Exhaustivity [Non_exhaustive]	0.99	0.92 – 1.07	0.856
Structure [baseline] × Exhaustivity [Non_exhaustive]	1.00	0.90 – 1.11	0.990
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	1.03	0.93 – 1.15	0.540
Random Effects			
σ^2	0.40		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.01		
$N_{Participants}$	36		
N_{Item}	720		
Observations	3897		
Marginal R^2 / Conditional R^2	0.000 / 0.015		

Table 26 – GLMM of fixations count in AOI6 per condition.

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.31	5.27 – 5.34	**<0.001**
Structure [baseline]	-0.03	-0.06 – 0.01	0.191
Structure [Only'-sentence]	-0.02	-0.05 – 0.02	0.427
Exhaustivity [Non_exhaustive]	0.01	-0.03 – 0.05	0.580
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.04	-0.01 – 0.10	0.145
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.01	-0.04 – 0.07	0.617
Random Effects			
σ^2	0.08		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.11		
$N_{Participants}$	36		
N_{Item}	704		
Observations	2763		
Marginal R^2 / Conditional R^2	0.003 / 0.117		

Table 27 – LMM of first fixation duration in AOI6.

4.2.6.4. Area of the critical segment (AOI8)

The sentence-final region (AOI8) was the critical segment for the stimuli, because this was the area where readers had their semantic expectations fulfilled or violated. As this was the last area, participants could not enter areas later than AOI8, so we did not analyse regressions-into AOI8, but regressions-out.

An initial analysis shows that the proportions of regressions out of AOI8 were lower in clefts than in other structures (baselines and 'only'-sentences):

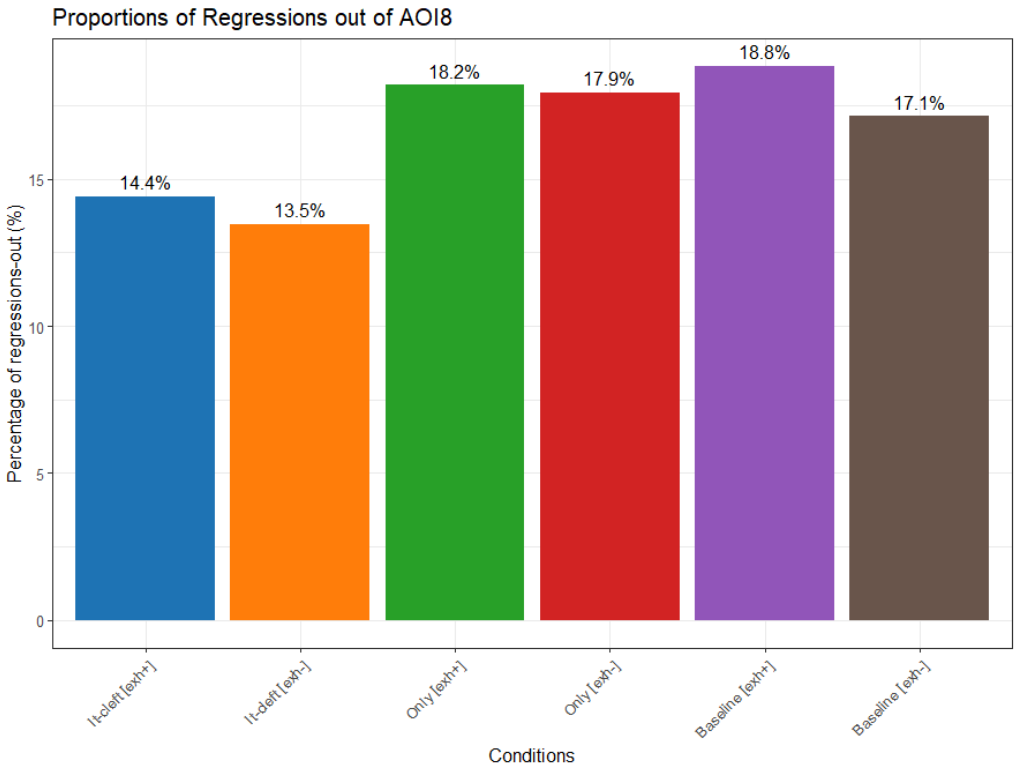


Figure 49 – Proportions of regressions out of AOI8 per condition.

A GLMM confirmed a processing advantage for clefts, as it revealed significant effects of baseline (IRR = 1.31, 95% CI [1.17, 1.47], $p < 0.001$) and 'only'-sentences (IRR = 1.27, 95% CI [1.13, 1.43], $p < 0.001$) (Table 28).

	Regression_out_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.67	0.58 – 0.76	**<0.001**
Structure [baseline]	1.31	1.17 – 1.47	**<0.001**
Structure [Only'-sentence]	1.27	1.13 – 1.43	**<0.001**
Exhaustivity [Non_exhaustive]	0.94	0.83 – 1.06	0.316
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.97	0.82 – 1.15	0.707
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	1.05	0.89 – 1.24	0.555
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Subject}$	0.10		
$N_{\ Subject}$	36		
$N_{\ Item}$	720		
Observations	4254		

Table 28 – Proportions of regressions out of AOI8 per condition.

Moreover, we looked into regression-path duration across conditions and observed that regressions in baselines were the longest, in contrast to the shortest regressions in clefts (Figure 47). Such findings match the results of the LMM on Table 28, which showed a positive effect for both structures (baselines and 'only'-sentences), but a negative effect for non-exhaustivity. In conclusion, 'only'-sentences and baselines made participants spend more time in regressions, whereas non-exhaustivity apparently facilitated the reading[33].

[33] This effect of facilitation - or acceptability, in the case of AJTs - has been quite unexpected since the first acceptability task that we designed. However, we will discuss in later sections that, for the eye-tracking tasks, one might make sense of such an effect. Considering that the items were presented after a sentence that mentioned a cooperative context, it would be only natural that readers would expect information about both subjects mentioned in the previous sentence.

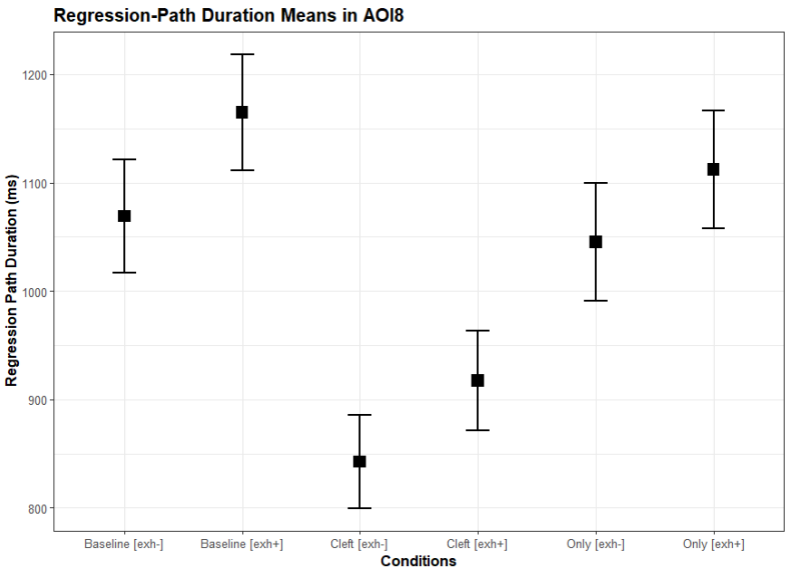


Figure 50 – Regression-path duration means in AOI8.

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.65	6.53 – 6.77	**<0.001**
Structure [baseline]	0.27	0.17 – 0.36	**<0.001**
Structure [Only'-sentence]	0.21	0.12 – 0.31	**<0.001**
Exhaustivity [Non_exhaustive]	-0.11	-0.21 – -0.01	**0.027**
Structure [baseline] × Exhaustivity [Non_exhaustive]	-0.01	-0.15 – 0.13	0.873
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	0.04	-0.10 – 0.18	0.581
Random Effects			
σ^2	0.44		
$\tau_{00\ Item}$	0.07		
$\tau_{00\ Participants}$	0.09		
ICC	0.27		
$N_{\ Participants}$	36		
$N_{\ Item}$	720		
Observations	4254		
Marginal R^2 / Conditional R^2	0.026 / 0.289		

Table 29 – LMM for regression-path duration in AOI8.

As we analysed fixation count and first-fixation duration, we noticed higher proportions of fixations into the critical segment (AOI8) of exhaustive baselines (Figure 50). This difference was later confirmed to be an effect of structure for the baselines (Table 30).

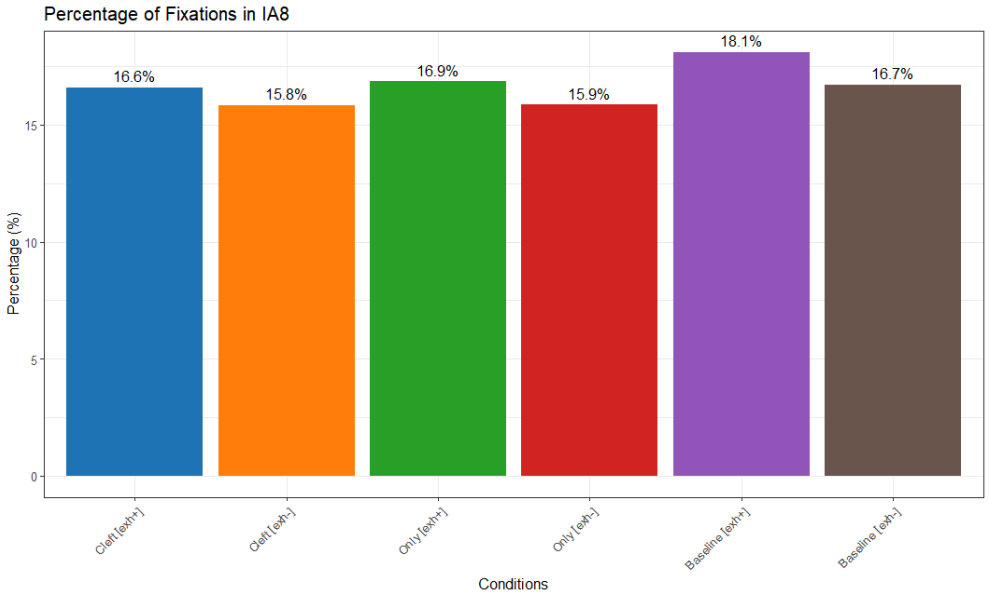


Figure 51 – Proportions of fixation count in AOI8 per condition.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	2.35	2.19 – 2.53	**<0.001**
Structure [baseline]	1.13	1.05 – 1.20	**0.001**
Structure [Only'-sentence]	1.05	0.98 – 1.12	0.179
Exhaustivity [Non_exhaustive]	0.94	0.88 – 1.01	0.074
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.96	0.87 – 1.06	0.404
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	1.01	0.91 – 1.11	0.904
Random Effects			
σ^2	0.34		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.03		
ICC	0.07		
$N_{Participants}$	36		
N_{Item}	720		
Observations	4056		
Marginal R^2 / Conditional R^2	0.008 / 0.080		

Table 30 – GLMM for fixation count in AOI8.

Interestingly, although baselines had significantly more fixations, first-fixation for these conditions had a negative effect on first-fixation duration (Table 31). This means that baselines had faster first-fixations than clefts. Non-exhaustivity also had a negative effect in this metric; in other words, non-exhaustive sentences had faster first-fixation durations.

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.46	5.41 – 5.51	**<0.001**
Structure [baseline]	-0.07	-0.11 – -0.04	**<0.001**
Structure [Only'-sentence]	-0.03	-0.07 – 0.00	0.088
Exhaustivity [Non_exhaustive]	-0.07	-0.11 – -0.03	**<0.001**
Structure [baseline] × Exhaustivity [Non_exhaustive]	0.02	-0.04 – 0.07	0.528
Structure [Only'-sentence] × Exhaustivity [Non_exhaustive]	-0.00	-0.05 – 0.05	0.975
Random Effects			
σ^2	0.10		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.02		
ICC	0.18		
$N_{\ Participants}$	36		
$N_{\ Item}$	720		
Observations	4057		
Marginal R^2 / Conditional R^2	0.014 / 0.189		

Table 31 – LMM for first fixation duration in AOI8.

4.2.7. Summary and Data Analysis

The data of this experiment was analysed across four areas of interest (AOIs), that target distinct stages of sentence processing: the focus marker (AOI2), the subject-NP(AOI3), the object-NP (AOI6), and the critical segment, which was also the final segment (AOI8). Altogether, the results reveal a non-uniform distribution of processing costs, with effects of structure and/or minor effects of exhaustivity.

In the focus marker region (AOI2): i) regression-in counts showed a positive effect of structure, but no effect of exhaustivity; ii) regression-path duration revealed a significant interaction between structure and non-exhaustivity. These results suggest that AOI2 - hence, focus markers - was a target to late integration processes.

The subject-NP region (AOI3) showed that: i) both structure and exhaustivity had effects on processing, as baselines and 'only'-sentences had more regressions-in than clefts; ii) 'only' and non-exhaustivity (combined) increased the likelihood of regressions, while non-exhaustivity alone decreased regressions; iii) regression-path duration showed a robust effect of structure, with baselines eliciting longer regressions than clefts, and a significant interaction between 'only' and non-exhaustivity, resulting in longer reinspection times; iv) 'only'-sentences received significantly more fixations and longer first fixation durations than clefts. In the object-NP area (AOI6), the regression-in counts showed a significant effect of structure, with baselines and 'only'-sentences increasing the likelihood of regressions.

Importantly, AOI3 reflects the locus where focus interpretation is computed and verified (Kiss, 1998; König, 1991; Lignani; Fonseca, 2023), while AOI6 represents the local integration point for the completion of the first event representation (Rayner, 1998; Lignani; Fonseca, 2023) - subject-NP + verb + object-NP. Increased regressions in 'only'-sentences and baselines indicate that readers rely on the subject-NP and on the object-NP to evaluate focus alternatives in 'only'-sentences or recover relevant information in the case of baselines. The early effects observed for 'only' could suggest that focus processing is initiated immediately upon encountering the subject-NP, rather than delayed to later integration stages, as an alternative to what may happen in clefts.

Finally, in the critical segment (AOI8): i) regression-out counts revealed a significant facilitation for clefts, which elicited fewer regressions out of the sentence-final region than baselines and 'only'-sentences; ii) fixation counts were significantly higher for baselines; iii) first-fixation duration showed a negative effect of both baselines and non-exhaustivity.

The results suggest that ‘only’-sentences and clefts differ not only in how they encode focus, but also in how they distribute processing costs across the sentence. As a focus-sensitive operator, ‘only’ triggers the immediate computation of alternatives and exhaustivity, leading to early processing costs in the subject region. This view of adverbs as focus markers aligns with Alternative Semantics (Rooth, 1992; Krifka, 1999, 2008). In contrast, clefts would allow smoother and more flexible initial processing, which might delay semantic evaluation to later regions of the sentence. This is consistent with previous research that argues that clefts would encode focus pragmatically (Horn, 1981; Kiss, 1998; Beaver; Clark, 2008). Such a pattern is consistent with models of incremental and predictive processing, in which different linguistic cues can modulate the timing of semantic integration (Pickering; Gambi, 2018), as well as with accounts of Good-enough processing, whereby initial interpretations may be underspecified and refined at later stages (Ferreira; Bailey; Ferraro, 2002).

4.3. Eye-tracking task in Brazilian Portuguese (BP)

After applying the eye-tracking task in English, we develop the same kind of experiment in BP. We made some adjustments in the experimental design, after having analysed the results of the previous task. The main features of the previous task (in English) were preserved: context-based items, ‘só’-sentences, clefts and exhaustivity/non-exhaustivity. We will discuss the changes in the next subsections.

4.3.1. Materials

The eye-tracking task in BP contained 32 experimental sets[34], plus 32 sets of fillers. We have also included 5 training sets before the task started. Each set was composed of two sentences: 1) a context-sentence, which necessarily mentioned two subject-NPs in a cooperative context; 2) a sentence that appeared in four different conditions. All the sentences were formed with transitive verbs, and their subject-NPs were carefully chosen so as half of the sets (16) presented female subject-NPs, and half of them presented male ones in order to minimize gender biases.

[34] See Appendix E for all the items.

The second sentences of the sets were the experimental items, and their 4 conditions were : a) '*só*'-sentences, conveying exhaustivity ([+exh]); b) '*só'*-sentences, but non-exhaustive ([-exh]); c) cleft, conveying exhaustivity ([+exh]); and d) cleft, but non-exhaustive ([-exh]). We can see an example of set in (14):

(14) Context-sentence: Ulisses e Evandro têm uma padaria no bairro deles.
Ulisses and Evandro own a bakery in their neighbourhood.

a. Hoje só o **Ulisses** assou os pães e os bolos também. ('*só*'; [+exh])
Today only Ulisses baked some bread and some cakes too.

b. Hoje só o **Ulisses** assou os pães e **o Evandro** também. ('*só*'; [-exh])
Today only Ulisses baked some bread and Evandro too.

c. Hoje foi o **Ulisses** quem assou os pães e os bolos também. (cleft; [+exh])
Today it was Ulisses who baked some bread and some cakes too.

d. Hoje foi o **Ulisses** quem assou os pães e o Evandro também. (cleft; [-exh])
Today it was Ulisses who baked some bread and Evandro too.

The 4-conditions design was the most significant change that we made in this experiment: while the eye-tracking task in English had experimental items across six conditions, we kept only four conditions for the BP task. We decided to remove the baseline conditions, as we understood that they were uninformative for the purposes of this experiment. Additionally, the baseline conditions could make the data analysis unnecessarily more complicated, as baseline-sentences hindered the crossing of the two independent variables that were present in the other items.

It is important to recall that the '…did too' structure used in the English version of the eye-tracking task differs from the '*e... também*' in BP, as the latter did not contain the auxiliary verb. For the BP items, we had no good alternatives to use VP-ellipsis instead of *gapping* structures, so we tried to address ambiguity biases by choosing verbs that could not be ambiguous about their object.

As we did with the English task, all the sentences (context-sentences and target-sentences) were presented in full on the screens. However, the segmentation of the target-sentences into areas of interest (AOI) for this experiment was slightly different than the previous one (Table 26).

Adverb (AOI1)	**Focus marker (AOI2)**	**Subject-NP (AOI3)**		**Verb (AOI4)**	**Object-NP (AOI5)**	**Critical segment (AOI6)**	**Closure (AOI7)**
Hoje Today	*só* only	*o Ulisses* Ulisses		*assou* baked	*os pães* some bread	*e os bolos* and some cakes	*também.* too.
Hoje Today	*só* only	*o Ulisses* Ulisses		*assou* baked	*os pães* some bread	*e o Evandro* and Evandro	*também.* too.
Hoje Today	*foi* it was	*o Ulisses* Ulisses	*quem* who	*assou* baked	*os pães* some bread	*e os bolos* and some cakes	*também.* too.
Hoje Today	*foi* it was	*o Ulisses* Ulisses	*quem* who	*assou* baked	*os pães* some bread	*e o Evandro* and Evandro	*também.* too.

Table 32 – Areas of interest (AOIs) of the eye-tracking experiment in BP.

We addressed a few issues with this new segmentation. Firstly, in our previous task (in English), the AOI that contained the critical segment matched the closure of the target-sentence, which we believed to be far from ideal. Because of that, in the BP experiment the critical segment of the sentence (AOI7) was separated from the closure (AOI8). Secondly, we ignored the word '*quem*' ('who') present in the clefts for the purpose of counting the AOIs. In the previous English task, this section corresponded to AOI4, in spite of not being analysed – this area ('*who*') contained no significant information for the objectives of these experiments. Because of that, there was a reduction of the total number of AOIs (seven) in the BP experiment, compared to the AOIs of the previous task (eight). Nonetheless, every correspondent AOI analysed in the previous experiment was analysed in this experiment too.

4.3.2. Method and Apparatus

This experiment was developed at the *Faculdade de Letras* (*School of Letras*) of the Federal University of Minas Gerais – UFMG (Belo Horizonte, Brazil). The research was approved by the *Comitê de Ética em Pesquisa* (*Ethics Committee in Research*) of the Federal University of Juiz de Fora – UFJF (Juiz de Fora, Brazil) (CAAE: 92040525.5.0000.5147).

We conducted the experiment using an EyeLink 1000 Plus system (SR Research Inc., Ontario, Canada), at a sampling rate of 1000 Hz. The experiment was designed on

Experiment Builder, with four versions (counterbalances 1, 2, 3, and 4), and the procedure was the same as the previous task. All stimuli were randomised on Experiment Builder, and counterbalanced across the four lists, following a Latin square design (Abbuhl; Gass; Mackey, 2013).

Once participants came to the laboratory, we first asked them to read and sign the *Termo de Consentimento Livre e Esclarecido* (consent form – Appendix G). After giving participants a brief explanation about the task and about the buttons that they should press, we started testing the set-up calibration before and after each sentence. We recalibrated the equipment whenever it was necessary throughout the session.

A binocular mount and a chin-and-forehead rest minimised head movement and ensured accurate gaze tracking. Participants sat approximately 65 cm from a BenQ monitor, where sentences appeared in black 'Courier New' font (14-point) on a light-gray background (Rayner; Pollatsek, 1989).

Each trial began with the short context-sentences exhibited on the screen, then participants pressed a keyboard key that would take them to the next screen, with the experimental items (target-sentences). Participants had a practice section with five sentences before the trial.

We included a yes/no comprehension question for every target-sentence (32) and for every filler (32). Participants should press different keys on the keyboard to answer *Sim* (*Yes*) or *Não* (*No*), and another key to read the next set of sentences. In this experiment, we did not include any breaks.

4.3.3. Independent Variables

Exactly like the previous eye-tracking task, our independent variables were: 1) (non-) exhaustivity; 2) syntactic structure; and 3) area of interest (AOI). We focused our analysis on the different AOIs and we had two conditions of exhaustivity (exhaustive/non-exhaustive) and two conditions of syntactic structure ('*só*'-sentences/clefts).

4.3.4. Dependent Variables

Having analysed the results of the previous eye-tracking task, we determined to look into the same parameters for this one, so we analysed: 1) regression-in count; 2)

regression-out count; and 3) regression-path duration; 4) fixation count; 5) first fixation duration. Our expectations for such measures followed the findings of the previous experiment: we expected to find patterns that coincide with semantic reanalysis and/or difficulties of integration for the non-exhaustive conditions, and significant effects of structure that matched the understanding that non-exhaustive 'só'-sentences were more difficult than non-exhaustive clefts.

4.3.5. Participants

For this experiment, we collected data from 37 (thirty-seven) undergraduate students of the *Faculdade de Letras* (School of *Letras*), at the Federal University of Minas Gerais (UFMG). We counterbalanced the results table and analysed data from 36 people – nine participants per the different versions of the experiment. All participants joined the task voluntarily, receiving no compensation, and they were native speakers of Brazilian Portuguese, without any kind of reading impairments, and with an age mean of 21.8 years old.

4.3.6. Results

We analysed the following results on *RStudio* (R Core Team, 2026).

Before starting the analysis, we balanced the results' table, which gave us a total of 36 participants – 9 participants per version (9x4= 36). To do so, we removed data of one participant whose registers have been predominantly inaccurate (92% of error). As there were Y/N-questions of comprehension after every item, we considered data from participants that were able to achieve a minimum 55% rate in accuracy (17/32 questions). The balanced results table contained 7,169 observations.

Once again, we used linear mixed-effects models (GLMMs) with a Poisson distribution (Bates *et al*., 2015) for discrete measures (regression-in count, regression-out count, first-fixation count) and linear mixed-effects models (LMMs) for continuous log-transformed duration-based measures (regression-path duration, first-fixation duration and dwell time). In both kinds of models – *glmer* and *lmer* functions – participants and items were included as random effects.

It is important to recall the segmentation into AOIs in this experiment: 1) adverb; 2) focus marker; 3) subject-NP; 4) verb; 5) object-NP; 6) critical segment; 7) closure

(‘*também*’). There were four different levels for conditions in AOI2, namely: exhaustive cleft (*intercept*), exhaustive ‘só’-sentence, non-exhaustive cleft, non-exhaustive ‘só’-sentence. These levels represented the different combinations for structure (cleft/‘only’) and exhaustivity (exhaustive/non-exhaustive).

The statistical models were first designed with independent variables (exhaustivity or structure), and we proceeded further investigations if we found it necessary, depending on the results of the previous model. The metrics investigated in each AOI followed the same logic of the previous eye-tracking experiment: the areas before the critical segment (AOI2, AOI3 and AOI5) were expected to show effects for a) regressions-in and b) regression-path duration, as these features might represent reanalysis. On the other hand, the critical segment (AOI6) – containing the satisfaction or violation of semantic exhaustivity – was believed to be relevant for a) regressions-out, b) regression-path duration, c) fixation count and d) first-fixation duration. These measures would presumably indicate difficulties during the processing of this area.

4.3.6.1. Area of the focus markers (AOI2)

The GLMM’s results revealed no significant influence from exhaustivity or structure, alone or combined, for regressions into AOI2 (Table 33). Variance was estimated for participants (n= 36; τ_{00} = 0.13), and for items (n= 128; τ_{00} = 0.08).

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.65	0.51 – 0.83	**0.001**
Structure [Só]	0.90	0.67 – 1.22	0.505
Exhaustivity [N_exhaustive]	0.89	0.67 – 1.19	0.435
Structure [Só] × Exhaustivity [N_exhaustive]	1.22	0.80 – 1.85	0.360
Random Effects			
σ^2	0.88		
$\tau_{00\ Item}$	0.15		
$\tau_{00\ Participants}$	0.16		
ICC	0.26		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	852		
Marginal R^2 / Conditional R^2	0.002 / 0.262		

Table 33 – GLMM for regression count into the focus marker area.

We used a LMM for regression-path durations (Table 34), and found no significant differences for this metric either – there was variance for participants (n= 36; $\tau_{00} = 0.02$), but not for items (n= 128; $\tau_{00} = 0$):

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.52	5.43 – 5.60	**<0.001**
Structure [Só]	-0.01	-0.12 – 0.09	0.806
Exhaustivity [N_exhaustive]	0.05	-0.05 – 0.15	0.320
Structure [Só] × Exhaustivity [N_exhaustive]	-0.01	-0.16 – 0.14	0.883
Random Effects			
σ^2	0.30		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.02		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	852		
Marginal R^2 / Conditional R^2	0.002 / NA		

Table 34 – LMM for regression-path duration in the focus marker area.

The fixation count (Figure 51) analysis revealed a significant decrease in fixations in '*só*', compared to clefts (Table 28). However, this might be a natural effect, due to the differences of length between the two structures – '*foi*' (three characters, plus spaces) in clefts, and 'só' (two characters, plus spaces) in '*só*'-sentences.

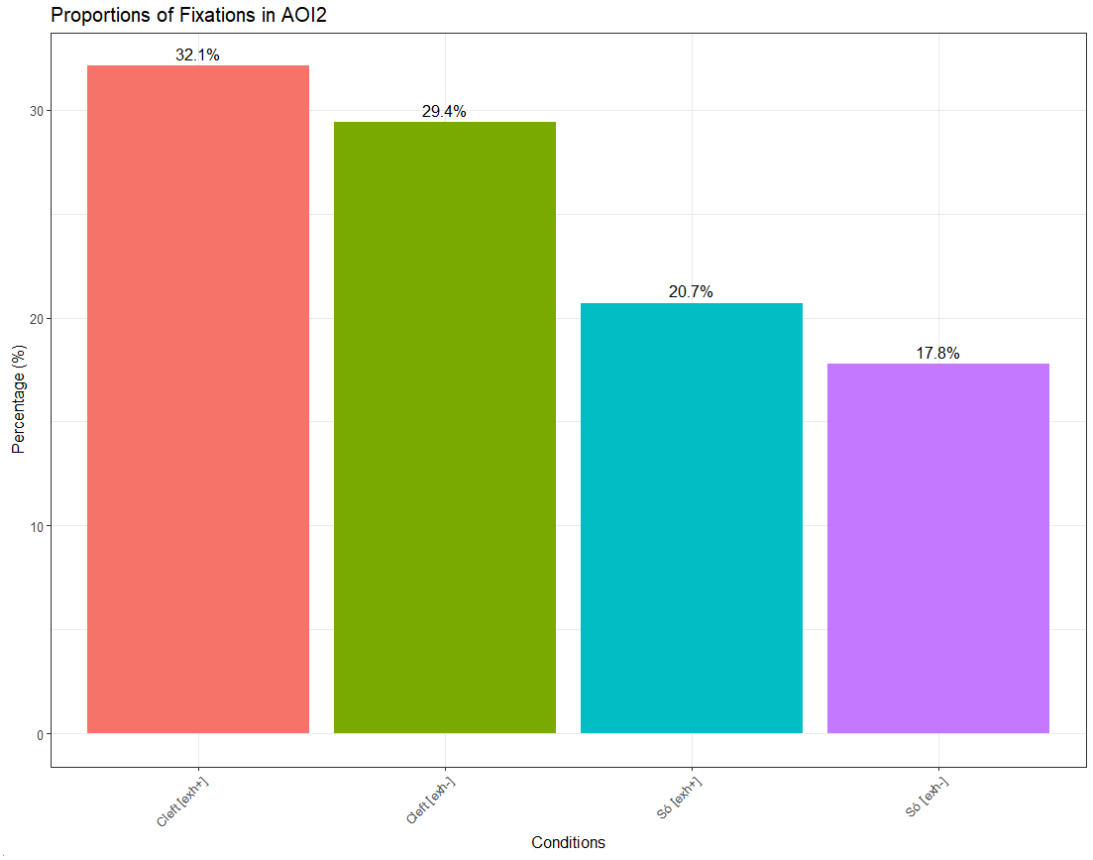


Figure 52 – Proportions of fixation count in the focus marker area (AOI2).

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	2.48	2.20 – 2.80	<**0.001**
Structure [Só]	0.82	0.68 – 0.98	**0.031**
Exhaustivity [N_exhaustive]	0.94	0.79 – 1.11	0.437
Structure [Só] × Exhaustivity [N_exhaustive]	1.13	0.87 – 1.48	0.356
Random Effects			
σ^2			
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
$N_{\ Participants}$	35		
$N_{\ Item}$	122		
Observations	404		

Table 35 – GLMM for fixation count in the focus marker area.

We designed another GLMM, comparing fixation counts across conditions in the AOI2 and found no evidence to believe that such effect of structure was not because of the differences in length (Table 36).

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	2.39	2.13 – 2.70	<**0.001**
Condition [Cleft [exh-]]	0.94	0.79 – 1.12	0.483
Condition [Só [exh-]]	0.86	0.71 – 1.05	0.149
Condition [Só [exh+]]	0.81	0.67 – 0.98	**0.027**
Random Effects			
σ^2	0.38		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.00		
ICC	0.00		
$N_{\ Participants}$	35		
$N_{\ Item}$	121		
Observations	389		
Marginal R^2 / Conditional R^2	0.018 / 0.018		

Table 36 – GLMM for fixation count in the focus marker area.

Besides, no effect of first-fixation duration was found for the AOI2, which supports the idea of no problems detected in this stage of the reading (Table 37).

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.19	5.07 – 5.31	**<0.001**
Structure [Só]	-0.11	-0.26 – 0.04	0.159
Exhaustivity [N_exhaustive]	-0.03	-0.17 – 0.11	0.691
Structure [Só] × Exhaustivity [N_exhaustive]	0.05	-0.17 – 0.27	0.653
Random Effects			
σ^2	0.29		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.03		
$N_{Participants}$	35		
N_{Item}	122		
Observations	412		
Marginal R^2 / Conditional R^2	0.006 / NA		

Table 37 – LMM for first fixation duration in the focus marker area.

4.3.6.2. Area of the subject-NP (AOI3)

This area contains the subject-NP (*e.g.*, Ulisses), that is, the scope of the focus markers. Overall, AOI3 showed more regressions-in for non-exhaustive 'só'-sentences, compared to their exhaustive counterparts or to either of the cleft conditions (Figure 52). These findings seemed to support our expectations about the non-exhaustive 'só'-sentences being the most difficult condition to process.

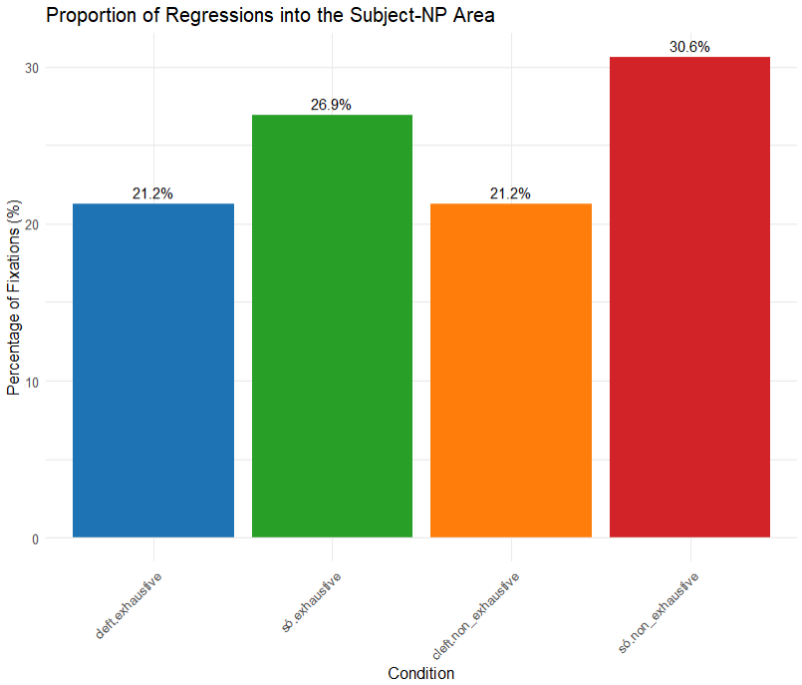


Figure 53 – Proportions of regressions into the subject-NP area.

Nevertheless, the GLMM (Table 38) shows that there is an effect of structure, but not for exhaustivity, concerning regressions into the AOI3:

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.44	0.34 – 0.57	**<0.001**
Exhaustivity [N_exhaustive]	0.98	0.75 – 1.29	0.913
Structure [Só]	1.42	1.09 – 1.83	**0.008**
Exhaustivity [N_exhaustive] × Structure [Só]	1.10	0.76 – 1.58	0.610
Random Effects			
σ^2	0.93		
$\tau_{00\ Item}$	0.09		
$\tau_{00\ Participants}$	0.27		
ICC	0.28		
$N_{Participants}$	36		
N_{Item}	128		
Observations	1096		
Marginal R^2 / Conditional R^2	0.030 / 0.299		

Table 38 – GLMM for regression-in count in the subject-NP area.

Such observation strengthens our argument that the semantic exhaustiveness has a milder effect on interpretation, when compared to syntactic structure, as there were no differences between exhaustive and non-exhaustive clefts.

No effects were found for regression-path duration in this area (Table 39):

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.02	5.89 – 6.15	**<0.001**
Structure [Só]	-0.04	-0.15 – 0.08	0.514
Exhaustivity [N_exhaustive]	0.03	-0.09 – 0.14	0.646
Structure [Só] × Exhaustivity [N_exhaustive]	0.00	-0.16 – 0.17	0.966
Random Effects			
σ^2	0.36		
$\tau_{00\ Item}$	0.01		
$\tau_{00\ Participants}$	0.10		
ICC	0.23		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	1060		
Marginal R^2 / Conditional R^2	0.001 / 0.230		

Table 39 – Regression-path duration means in the subject-NP area.

Although there was an increase of regressions into the subject-NP area (AOI3), that increase did not correspond with significant increases in regression-path duration times. This suggests that the regressions into this area were rapid confirmatory checks, to verify referential information, without substantial processing difficulties.

A descriptive analysis showed that the distribution of fixation count across conditions was quite homogeneous (Figure 53), and this was later confirmed by a GLMM that presented no main effects of structure or exhaustivity (Table 40).

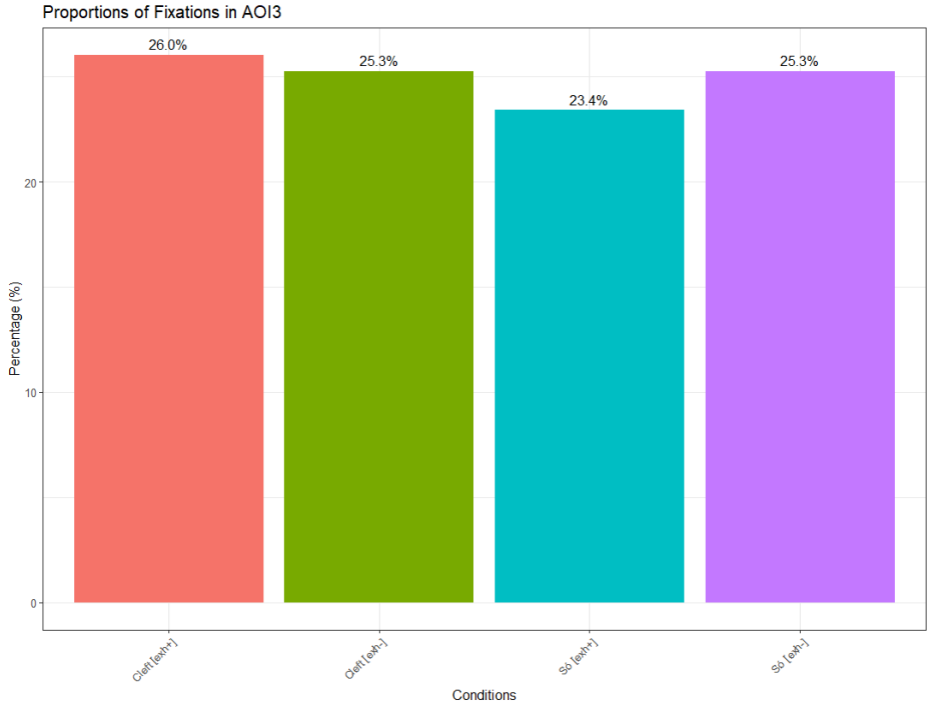


Figure 54 – Proportions of fixations into the subject-NP area.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	4.19	3.74 – 4.71	**<0.001**
Structure [Só]	0.99	0.91 – 1.09	0.891
Exhaustivity [N_exhaustive]	0.97	0.89 – 1.07	0.575
Structure [Só] × Exhaustivity [N_exhaustive]	1.04	0.92 – 1.18	0.533
Random Effects			
σ^2	0.21		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.09		
ICC	0.30		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	935		
Marginal R^2 / Conditional R^2	0.001 / 0.299		

Table 40 – GLMM for fixation count in the subject-NP area.

Moreover, no effects of first fixation duration were found either (Table 35).

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.24	5.14 – 5.35	**<0.001**
Structure [Só]	0.01	-0.13 – 0.14	0.924
Exhaustivity [N_exhaustive]	-0.01	-0.14 – 0.12	0.888
Structure [Só] × Exhaustivity [N_exhaustive]	-0.03	-0.22 – 0.17	0.790
Random Effects			
σ^2	0.28		
$\tau_{00\ Item}$	0.04		
$\tau_{00\ Participants}$	0.01		
ICC	0.15		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	934		
Marginal R^2 / Conditional R^2	0.001 / 0.151		

Table 41 – LMM for first fixation duration in the subject-NP area.

4.3.6.3. Area of the object-NP (AOI5)

The area that contained the object-NP of the sentence verb was the area where the readers found information about one of the people previously cited in the context-sentence. Hence, by the time participants have AOI5 read, they would probably expect information from the other person for the next segments. However, exhaustive conditions preserved semantic exhaustivity but failed to recall both people mentioned in the context-sentence.

A descriptive analysis showed the distribution of regressions into AOI5, showing that exhaustive 'only'-sentences received the highest percentage of regressions, while exhaustive clefts received the lowest percentage (Figure 54):

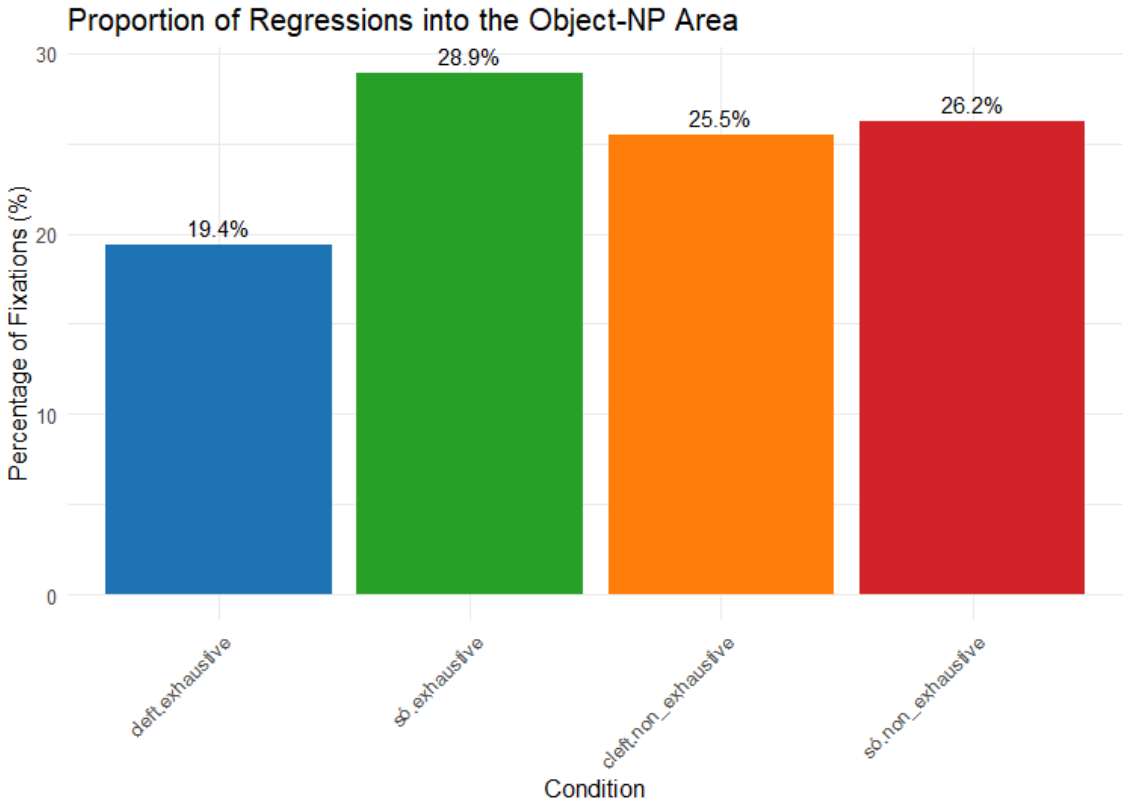


Figure 55 – Proportions of regressions into the object-NP area.

GLMMs revealed a significant increase of regressions into the AOI5 because of structure and a marginally significant increase because of exhaustivity as well, when they are isolated. Interestingly, when these variables are combined, they produce a marginally significant decrease of regressions into the object-NP area (Table 42).

	Regression_in_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.36	0.29 – 0.46	**<0.001**
Exhaustivity [N_exhaustive]	1.33	1.01 – 1.74	**0.043**
Structure [Só]	1.47	1.12 – 1.92	**0.005**
Exhaustivity [N_exhaustive] × Structure [Só]	0.68	0.47 – 0.99	**0.044**
Random Effects			
σ^2	1.08		
$\tau_{00\ Item}$	0.06		
$\tau_{00\ Participants}$	0.14		
ICC	0.15		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	1083		
Marginal R^2 / Conditional R^2	0.016 / 0.168		

Table 42 – GLMM for regression-in count in the object-NP area.

Although there was a significant effect for regressions into AOI5, like AOI3, there was no effect of regression-path duration (Table 43):

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.48	6.23 – 6.73	**<0.001**
Structure [Só]	-0.19	-0.52 – 0.13	0.246
Exhaustivity [N_exhaustive]	-0.00	-0.33 – 0.32	0.978
Structure [Só] × Exhaustivity [N_exhaustive]	0.01	-0.45 – 0.47	0.965
Random Effects			
σ^2	0.61		
$\tau_{00\ Item}$	0.36		
$\tau_{00\ Participants}$	0.09		
ICC	0.43		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	1083		
Marginal R^2 / Conditional R^2	0.008 / 0.433		

Table 43 – LMM for regression-path duration in the object-NP area.

We found no effects for fixation count (Figure 55; Table 44) or first fixation duration (Table 45) in AOI5.

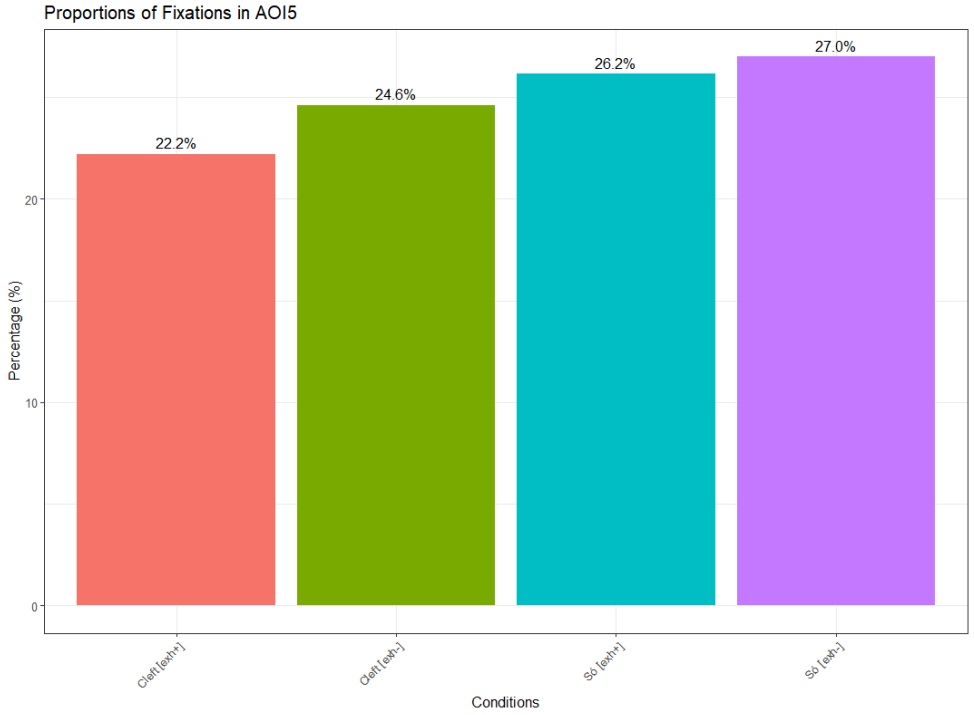


Figure 56 – Proportions of fixation count in the object-NP area.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	3.73	3.28 – 4.23	**<0.001**
Structure [Só]	1.15	0.99 – 1.32	0.064
Exhaustivity [N_exhaustive]	1.09	0.94 – 1.25	0.268
Structure [Só] × Exhaustivity [N_exhaustive]	0.90	0.74 – 1.11	0.331
Random Effects			
σ^2	0.21		
$\tau_{00\ Item}$	0.05		
$\tau_{00\ Participants}$	0.05		
ICC	0.33		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	957		
Marginal R^2 / Conditional R^2	0.009 / 0.332		

Table 44 – GLMM for fixation count in the object-NP area.

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.33	5.26 – 5.40	**<0.001**
Structure [Só]	-0.02	-0.10 – 0.06	0.685
Exhaustivity [N_exhaustive]	0.03	-0.05 – 0.11	0.479
Structure [Só] × Exhaustivity [N_exhaustive]	-0.00	-0.11 – 0.11	0.999
Random Effects			
σ^2	0.16		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.01		
ICC	0.11		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	953		
Marginal R^2 / Conditional R^2	0.002 / 0.109		

Table 45 – LMM for first fixation duration in the object-NP area.

4.3.6.4. Area of the critical segment (AOI6)

For the sentence critical region (AOI6) we analysed regressions-out instead of regressions-in, for the same reasons we did so in the previous task.

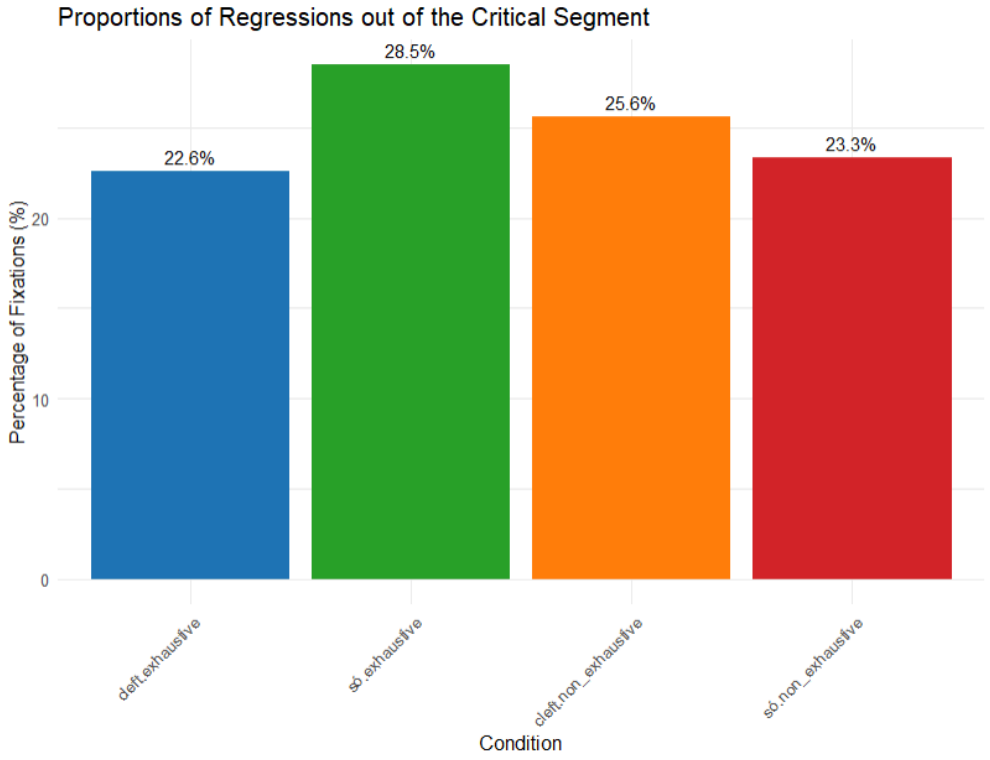


Figure 57 – Proportions of regressions out of the critical segment.

We designed a GLMM including the interaction between the two independent variables (structure and exhaustivity), illustrated in Table 37. However, this model produced warnings indicating that it was nearly unidentifiable. The estimated standard errors were extremely high, and the model diagnostics suggested numerical instability, indicating that the data did not support reliable estimation of the interaction term. Therefore, we designed another model for regressions-out, with no interaction between the variables. This model showed no significant effects from either (Table 46).

	Regression_out_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.46	0.46 – 0.46	**<0.001**
Exhaustivity [N_exhaustive]	1.15	1.15 – 1.16	**<0.001**
Structure [Só]	1.29	1.29 – 1.29	**<0.001**
Exhaustivity [N_exhaustive] × Structure [Só]	0.74	0.74 – 0.74	**<0.001**
Random Effects			
σ^2	0.94		
$\tau_{00\ Item}$	0.35		
$\tau_{00\ Participants}$	0.07		
ICC	0.31		
$N_{Participants}$	36		
N_{Item}	128		
Observations	1079		
Marginal R^2 / Conditional R^2	0.006 / 0.314		

Table 46 – GLMM with interaction for regressions out of the critical segment.

Likewise, regression-path duration showed no significant effects (Table 47):

	log(Regression_path_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.83	6.58 – 7.08	**<0.001**
Structure [Só]	0.17	-0.16 – 0.50	0.308
Exhaustivity [N_exhaustive]	0.03	-0.30 – 0.36	0.865
Structure [Só] × Exhaustivity [N_exhaustive]	-0.27	-0.74 – 0.20	0.257
Random Effects			
σ^2	0.71		
$\tau_{00\ Item}$	0.37		
$\tau_{00\ Participants}$	0.07		
ICC	0.38		
$N_{Participants}$	36		
N_{Item}	128		
Observations	1076		
Marginal R^2 / Conditional R^2	0.007 / 0.389		

Table 47 – LMM for regression-path duration in the critical segment.

Although there were no effects for regressions or regression-path duration, another GLMM later showed that the presence of the adverb in '*só*'-sentences conditions influenced the increase of fixation count.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	3.73	3.32 – 4.20	**<0.001**
Structure [Só]	1.15	1.01 – 1.30	**0.036**
Exhaustivity [N_exhaustive]	1.03	0.90 – 1.17	0.683
Structure [Só] × Exhaustivity [N_exhaustive]	1.02	0.85 – 1.22	0.851
Random Effects			
σ^2	0.21		
$\tau_{00\ Item}$	0.04		
$\tau_{00\ Participants}$	0.05		
ICC	0.29		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	1043		
Marginal R^2 / Conditional R^2	0.018 / 0.303		

Table 48 – GLMM for fixation count in the critical segment.

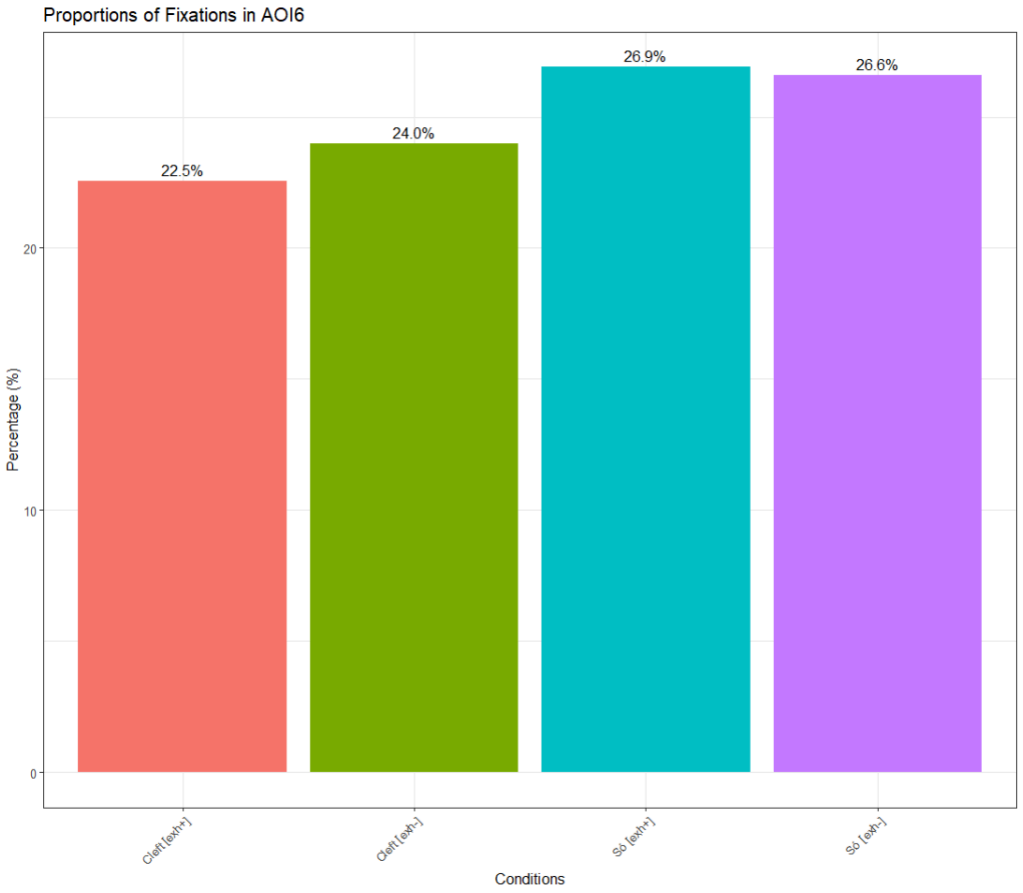


Figure 58 – Proportions of fixation count in the critical segment.

In order to conduct a deeper analysis about the fixation count, we fitted a GLMM with the conditions as variables. Results show that the non-exhaustive '*só*'-sentences had significantly more fixations than non-exhaustive clefts (Table 49), which supports our expectations about how violations of exhaustivity in sentences with the adverb would be more difficult to process than violations in clefts. Differently from the focus marker area (AOI2), the critical segment (AOI6) was the same for all conditions, so we have no reason to believe that this contrast might be due to differences in the length of the AOI6 in each syntactic structure.

	Fixation_Count		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	3.73	3.32 – 4.20	**<0.001**
Condition [Cleft [exh-]]	1.03	0.90 – 1.17	0.683
Condition [Só [exh-]]	1.20	1.05 – 1.36	**0.006**
Condition [Só [exh+]]	1.15	1.01 – 1.30	**0.036**
Random Effects			
σ^2	0.21		
$\tau_{00\ Item}$	0.04		
$\tau_{00\ Participants}$	0.05		
ICC	0.29		
$N_{\ Participants}$	36		
$N_{\ Item}$	128		
Observations	1043		
Marginal R^2 / Conditional R^2	0.018 / 0.303		

Table 49 – GLMM for fixation count comparing conditions in the critical segment.

Despite the differences in fixation count, the LMM for regression-path duration showed no effects for exhaustivity or structure (Table 50).

	log(First_fixation_Duration)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.39	5.32 – 5.45	**<0.001**
Structure [Só]	-0.02	-0.09 – 0.05	0.652
Exhaustivity [N_exhaustive]	-0.03	-0.10 – 0.04	0.453
Structure [Só] × Exhaustivity [N_exhaustive]	0.03	-0.07 – 0.13	0.546
Random Effects			
σ^2	0.16		
$\tau_{00\ Item}$	0.00		
$\tau_{00\ Participants}$	0.02		
$N_{Participants}$	36		
N_{Item}	128		
Observations	1022		
Marginal R^2 / Conditional R^2	0.001 / NA		

Table 50 – LMM for first fixation duration in the critical segment.

4.3.7. Summary and Data Analysis

In the focus marker region (AOI2), we found: i) no regression-based effects, which could be evidence that readers did not engage in semantic or structural reanalysis in this early area; ii) a significant decrease in fixation count for '*só*'-sentences, which might be due to differences on the length of the words presented in this area ('*só*' vs. '*foi*'), rather than by processing differences related to focus interpretation.

In the area of the subject-NP (AOI3), there was an increase of regressions-in for sentences with '*só*'. This means that readers were more likely to revisit this region during the processing of '*só*'-sentences, even though the absence of corresponding effects in regression-path duration suggests that these regressions were not associated with substantial processing difficulty. This dissociation points to a pattern of rapid, low-cost reinspection once participants had to reanalyse the sentences. For the area of the object-NP (AOI5): i) 's*ó*'-sentences had significantly more regressions-in than clefts; ii) there was a main effect of structure, a main effect of (non-)exhaustivity and a main effect of these variables combined.

The critical region (AOI6) provides the clearest evidence of processing consequences associated with exhaustivity violations, as we observed an increase in fixation counts for non-exhaustive '*só*'-sentences indicate local processing effort, suggesting that readers are sensitive to the inconsistency introduced at this point. However, the absence of effects in regression-out and regression-path duration suggests that this difficulty did not trigger systematic reanalysis. Instead, readers appear to register the anomaly without fully revising their interpretation. There is evidence to believe that this might be due to the fact that the disruption of semantic exhaustivity was introduced by mentioning one of the names introduced in the context-sentence, which possibly softens the violation of exhaustivity (Krifka, 2008).

Taken together, the results support a model of sentence processing in which readers rely on incremental, expectation-driven mechanisms that do not always result in full semantic computation. While focus marking influences early structural processing, the semantic consequences of exhaustivity are only partially integrated and often do not trigger complete reanalysis.

In particular, the dissociation between increased fixation counts and the absence of regression-based effects suggests that readers may detect inconsistencies without engaging in effortful reinterpretation. This pattern aligns with Good-Enough approaches to language processing (Ferreira, 2018), according to which comprehenders construct shallow representations and may fail to fully resolve semantic conflicts unless required to do so.

Furthermore, the stronger effects observed in '*só'*-sentences compared to clefts suggest that lexical focus markers provide more immediate cues for expectation formation, leading to more pronounced local disruption when those expectations are violated. In contrast, clefts may allow for more flexible interpretation, reducing the need for immediate processing adjustments.

The results reveal a consistent dissociation between early structural processing and later interpretive mechanisms. Apparently, semantic exhaustivity did not uniformly drive processing difficulty, suggesting that its computation might be delayed or even optional, in some cases. On the other hand, the effect of structure was considerable in practically every AOI, suggesting that there are indeed semantic differences between clefts and 'só'-sentences. Importantly, even though AOI2 was the area of the focus marker, AOI3 is the area that contains the scope of such marker – the name of the person who is on focus. Such

interpretation might suggest that the focus marker 'attracted' attention to its scope (Kiss, 1998) when readers found any kind of difficulties. The fact that '*só*'-sentences marked their scope in a stronger way than clefts is consistent both with the theoretical background and the previous findings in this research.

4.4. Data Comparison and Analysis for Eye-Tracking Tasks

The comparison between the English and the BP eye-tracking experiments reveals both converging and diverging patterns for these languages in what concerns the processing of focus and exhaustivity. In both languages, the results suggest that the computation of focus – therefore, of its semantic effects – is not uniformly immediate, but emerges at different stages of processing depending on the structure.

In English, 'only'-sentences showed clear effects compared to clefts, in the focus markers area (AOI2) and in the subject-NP (AOI3), representing the areas where the focus particles and their scope were, respectively. These effects include early measures (longer first fixation duration; more fixation count) and late measures (increased regressions-in; longer path-duration). Altogether, these might be evidence that readers engage in immediate computation of focus alternatives and exhaustivity, according to Alternative Semantics (Rooth, 1992; Krifka, 1999, 2008). Conversely, clefts exhibited fewer regressions into earlier regions such AOI2 or AOI3, but longer first fixation durations in the critical segment (AOI8). This suggests a redistribution of processing cost, whereby structural encoding of focus allows for smoother initial parsing but delays the evaluation of exhaustivity until sentence-final integration (Kiss, 1998; Beaver; Clark, 2008; Drenhaus, Zimmerman, Vasishth, 2011).

On the other hand, in Brazilian Portuguese (BP) this pattern of more complexity was less robust and more delayed: although '*só*'-sentences showed increased regressions in the subject-NP region (AOI3), these were not accompanied by longer regression-path durations or early fixation effects, which might suggest rapid confirmatory checks rather instead of difficult integration. Crucially, effects of exhaustivity emerged more clearly in the critical segment (AOI6), where non-exhaustive '*só*'-sentences elicited significantly more fixations than clefts. This suggests that the processing cost associated with exhaustivity violations in BP is delayed and only becomes detectable once sufficient propositional information is

available for integration, consistent with incremental and Good-enough processing accounts (Ferreira; Bailey; Ferraro, 2002).

We argue that these cross-linguistic differences do not depend exclusively on the lexical properties of the adverbs only and '*só*' – that are quite similar in both languages – but on the differences between the syntactic structures that function as items in each experiment. The English stimuli used VP-ellipsis ('…did too'), which creates a tightly constrained contrast between parallel predicates, and strengthens expectations about exhaustivity and facilitates the detection of mismatches. As a result, when these expectations are violated, readers are more likely to engage in repeated regressions to resolve the conflict between the encoded meaning of 'only' and the information provided in the second conjunct. Conversely, the BP stimuli relied on *gapping* ('e o … *também*'), a less explicit predicate repetition strategy that requires structural reconstruction and parallelism (Johnson, 2009). This might weaken the contrast and allow for a more flexible interpretation.

Consequently, although '*só*' still encodes exhaustivity, violations do not trigger the same degree of reanalysis, resulting in fewer or less systematic regression patterns. This interpretation is consistent with accounts of incremental and Good-enough processing, according to which readers may initially adopt underspecified representations and only engage in deeper reanalysis when structural and semantic cues strongly support it (Ferreira; Bailey; Ferraro, 2002).

It is difficult to determine our expectations for the cross-linguistic comparisons if the stimuli used in both eye-tracking tasks followed the same syntactic pattern, for example, constituting a *gapping* sentence in each and every case. However, the similar effects observed for the AJTs 2 and 3, that is, the Likert tests with *gapping* items in English (AJT 2) and in BP (AJT3), may suggest that there could have been similar results for the eye-tracking tasks too.

5. METHODOLOGICAL DISCUSSIONS AND FURTHER DIRECTIONS

Despite the contributions of the present study, some methodological and design-related limitations should be acknowledged. A central issue concerns the structural differences between the English and BP stimuli. Two out of three English experiments used VP-ellipsis ('…did too') structures, whereas the BP tasks relied on *gapping* ('e *também…'/'e… também'*), which involves more complex structural reconstruction and less explicit predicate repetition (Johnson, 2009). As previously discussed, this difference may have influenced the results, particularly in the judgment of acceptability for AJTs 1 and 2, and late measures (such as regressions) in the eye-tracking experiments.

On the one hand, this methodological choice allowed an indirect comparison between these two types of structures, concerning focus and exhaustivity processing; on the other hand, cross-linguistic comparisons should be interpreted with caution, as some of the differences may reflect structural variation rather than purely lexical or semantic contrasts between 'only' and '*só*' (Carlson, 2001; Clifton; Bock; Rado, 2000; Johnson, 2009). In fact, there is evidence to believe that 'only' and '*só*' function in the exact same way as focus particles (König, 1991; Kiss, 1998; Carlson, 2014; Lignani; Fonseca, 2023). That is to highlight that the cross-linguistic comparisons should not be interpreted as a result of semantic properties of the adverbs 'only' and 'só' purely. Instead, it is important to recognise methodological limitations and interpret the results considering the interaction between those focus particles and the whole syntactic structure where they are inserted in.

Concerning the Likert tests (Likert, 1932) developed in this research, the so-called 'AJT's, there are a few considerations to be made. Firstly, this experiment was online and remotely conducted, so we could not observe and/or use strategies to interfere in participants' focus towards the task. Because of that, we inserted ungrammatical fillers to use them as attention-filters afterwards, removing from the analyses participants who failed to classify most sentences (80%) as unacceptable – this was discussed in the Methodology section. Ideally, this strategy would be a good one to make sure that the data analysed was from attentive participants, hence to increase the reliability of the results. However, we reevaluate this as a potential problem for the very ranking of sentences, which was the task itself for this experiment: as we included ungrammatical items, with at least one clear syntactic and/or semantic problem, such items would lower the means of acceptability for the experimental items. In other words, since participants had to deal with indisputable unacceptable items,

they could eventually become more likely to rate grammatical items as acceptable ones, regardless of semantic problems (such as non-exhaustivity).

Another matter that is worthy to address is the Likert scale. There are some possible variations for this scale, and each one of them has different strengths and issues (Sá; Ciríaco; Godoy, 2022). Our choice was a 5-point-scale because: i) it is an odd-numbered (symmetric) scale, so it has a midpoint that represents indeterminate judgments (neutral) for a task involving subtle semantic contrasts, as exhaustivity processing; ii) in comparison to shorter scales (a 3-point-scale), it is better to preserve the variability in the responses; iii) compared to longer scales (*e.g.*: a 7-point-scale), it is best in avoiding ambiguity and uneven response distributions. Moreover, from a statistical point of view, 5-point-scales suit particularly well ordinal modeling (such as CLMMs), given that they contribute for a more balanced distribution, thus, data density, across their different categories. By doing so, this type of scales reduce sparsity and contribute to more stable parameter estimation (Josh *et al.*, 2015).

One last consideration on AJTs concerns the variation between participants. Participated in the tasks in English (AJTs 1 and 2) people with the same age mean (≈19 years old) and the same scholarly levels, as all of them were students in the School of Psychology at the University of Leicester. Differently, the Likert test (AJT3) developed in BP was completed by very different profiles of participants, with different education (17 undergraduates, 8 graduated and 15 postgraduates) and different ages, with a mean of approximately 28 years old. It is important to consider these differences, as they may have had an impact on the differences of the results. This disparity among participants from the English AJTs and the BP one was not a methodological choice - in fact, it was merely circumstantial and it was impossible to control.

Regarding the eye-tracking tasks, there are also a few issues to address. An important difference concerns the asymmetry in the experimental design across languages. The eye-tracking task in English included three conditions of structure: two focus-marked sentences (‘only’-sentences and it-clefts), plus two baseline conditions, that is, sentences with no ‘only’ and no cleft. Crossing those variations with the two conditions of exhaustivity (exhaustive/non-exhaustive), there were a total of six conditions. Differently, the eye-tracking experiment in BP presented a total of four conditions: exhaustive/non-exhaustive crossed with two conditions of structure, namely ‘*só*’-sentences and *clivadas* (clefts). This difference, that

is, the absence of baseline conditions in the BP task, has direct implications for cross-linguistic comparisons.

When we designed the task in English, we expected the baseline conditions to provide a neutral point of reference, allowing us to disentangle effects associated specifically with focus-marking strategies from more general processing patterns. However, we later realised that, for the purposes of this research, it would be best to compare focus-marked sentences between one another, rather than using a neutral sentence previously. For this reason, we decided to remove baseline conditions for the BP task, making statistical analyses more accurate. To minimize problems of comparison between the English eye-tracking task and the BP one, we used the same intercept in all analyses (exhaustive clefts), so that the effect observed would be comparable cross-linguistically. Despite that, it is important to consider the fact that the English design had two conditions more than the BP one, thus preventing the results of these tasks to be perfectly comparable among one another. Additionally, there was also a difference between the age means for the participants of the task in English (μ= 19 years old), and for people who participated in the task in BP ($\mu \approx$ 22 years old).

Finally, another limitation concerns the interpretation of eye-tracking measures. First fixation duration, fixation count and regression-based metrics provide valuable insights into processing dynamics, but they cannot map directly onto specific cognitive operations. As such, interpretations regarding the timing of exhaustivity computation should be understood as inferential rather than definitive. Further directions for this research involve complementary methodologies, such as the electroencephalography (EEG) technique, could help precise exactly how, when and where the semantic processing occurs for these types of sentences.

6. LAST CONSIDERATIONS

The present study contributes to an experimental perspective of semantics and pragmatics towards the processing of semantic exhaustivity. It is largely discussed how exhaustivity is not processed uniformly across linguistic constructions; however, there is certain consensus that it depends on how it is encoded – lexically (syntactically) and/or discourse-pragmatically. This thesis aimed to propose a unified account of how exhaustivity is predicted and integrated in 'only'-sentences and in clefts, based in cross-linguistic evidence from English and Brazilian Portuguese, literature accounts and empirical evidence from the tasks developed: the acceptability judgment tasks (AJTs) and eye-tracking experiments.

In 'only'-sentences, exhaustivity is lexically encoded: the focus particle directly introduces a set of alternatives and excludes them (Rooth, 1992; Krifka, 2008; Lignani; Fonseca, 2023). This encoding leads to strong and immediate interpretive commitments. In contrast, clefts do not encode exhaustivity lexically. Rather, exhaustivity emerges through discourse structure and pragmatic inference (Horn, 1981; Velleman *et al.*, 2012; Pollard; Yasavul, 2016; Onea, 2019). According to Onea (2019), clefts presuppose a question under discussion (QUD) and provide an answer interpreted as maximally informative. Thus, exhaustivity is processed as an inference derived from the relation between the utterance and the relevant set of alternatives, rather than as a fixed semantic property (Kiss, 1998; Beaver; Clark, 2008).

This distinction suggests that 'only'-sentences should trigger stronger and earlier expectations of exclusivity, whereas clefts should allow for more flexible and context-dependent interpretations, which is consistent with the results of the empirical core of this research.

Across the AJTs, clefts were consistently rated as more acceptable than 'only'-sentences, particularly than non-exhaustive ones. This suggests that clefts allow for more flexible interpretations of exhaustivity, which is consistent with their pragmatic derivation. Importantly, the type of additive structure modulated these effects. In English (AJT1), where VP-ellipsis was used, clearer interactions between structure and exhaustivity emerged, reflecting stronger constraints on interpretation. In contrast, in tasks using *gapping* structures (AJT2 and AJT3), the effects were attenuated, suggesting that structural complexity can modulate the salience of exhaustivity violations.

Likewise, the eye-tracking experiment in English revealed a clear dissociation between ‘only’-sentences and clefts: ‘only’-sentences elicited early effects in the focus marker and subject regions (AOI2 and AOI3), including increased fixations and regressions; clefts showed fewer early disruptions, but greater processing cost in the critical region (AOI8), particularly in first-fixation duration. This pattern indicates that ‘only’-sentences trigger early predictive commitments, whereas clefts delay the computation of exhaustivity until later integration stages. Additionally, shorter regression-path durations in ‘only’-sentences, compared to clefts, suggest that these structures may be rejected or reinterpreted more rapidly; conversely, clefts appear to require more sustained integration, potentially due to their structural complexity, length, and greater sensitivity to contextual influences.

As discussed in previous sections, the eye-tracking in BP, showed a more attenuated and delayed pattern: effects in early regions (AOI2, AOI3) were weaker and often limited to regression counts without corresponding increases in duration. Crucially, effects emerged more clearly in the critical segment (AOI6), where non-exhaustive *só*-sentences elicited more fixations than clefts. This might be evidence to the fact that the computation of exhaustivity in BP is delayed and less robust, not because ‘*só*’ differs fundamentally from ‘only’, but because the use of *gapping* reduces the salience of contrasts and shifts processing demands toward later integration[35].

This asymmetry suggests that comprehensioners initially commit to an exclusive interpretation in ‘only’-sentences and must subsequently revise it when confronted with contradictory information, resulting in increased regressions. In this respect, the processing of exhaustivity violations in ‘only’-sentences closely resembles garden-path effects: readers are led toward a strongly preferred interpretation and are forced to backtrack when incoming input conflicts with their expectations (Drenhaus; Zimmerman; Vasishth, 2011; Zimmerman, 2020).

On the other hand, clefts allow for underspecified representations that can be incrementally refined. Consequently, violations of exhaustivity in clefts are less disruptive and more dependent on later stages of integration, supporting accounts in which their meaning emerges from the interaction between focus, context, and pragmatic reasoning. Most experimental research suggests that the interpretation of exhaustivity in clefts is pragmatical, not semantically encoded (Destruel *et al*., 2015; Pollard; Yasavel, 2016; Onea, 2019). Because

[35] Importantly, the cross-linguistic comparison further shows that these predictive processes are modulated by syntactic structure.

of that, exhaustivity may be inferred rather than fully computed, allowing readers to maintain a flexible representation that can accommodate subsequent *input* with minimal disruption.

The Good-Enough approach (Ferreira, Bailey; Ferraro, 2002), in Psycholinguistics, offers a possible support for this account: readers may initially construct shallow or underspecified representations of exhaustivity, especially in structures where the cue is less explicit or more context-dependent, as in clefts. However, once they encounter a problem in a sentence - or processing difficulties - they would reanalyse it.

Crucially, the present findings can be further understood in terms of predictive processing mechanisms. Within a predictive framework, the observed differences between constructions can be further explained by distinguishing between two types of predictive mechanisms (Pickering; Gambi, 2018): i) prediction-by-association: a fast, automatic process driven by lexical cues and semantic/syntactic activation; ii) prediction-by-production: a slower, inferential process in which comprehenders simulate the speaker's intentions and reconstruct the communicative context.

We propose that 'only'-sentences primarily engage *prediction-by-association*. The presence of the focus particle immediately activates alternatives and generates strong expectations of exclusivity, as the exhaustivity, in this case, is processed in-locus. This leads to early processing effects – such as we have observed for the eye-tracking task. Clefts, by contrast, rely more heavily on *prediction-by-production.* Because exhaustivity is not lexically encoded, comprehenders must infer the intended meaning by reconstructing the QUD and the speaker's communicative goals. This results in weaker, more gradual predictions that emerge through integration.

Taken together, these findings support a model in which the prediction of exhaustivity depends on both linguistic encoding and structural context.

1. In 'only'-sentences, exhaustivity is lexically encoded, triggering early, strong predictions through prediction-by-association.

2. In clefts, exhaustivity is pragmatically derived, emerging through prediction-by-production and discourse integration.

3. Structural differences, such as VP-ellipsis versus *gapping*, modulate the timing and strength of mismatch detection, particularly in later processing stages.

To account for these patterns, we propose that exhaustivity is not computed through a single unified mechanism, but rather through two qualitatively distinct predictive routes. In ‘only’-sentences, exhaustivity is triggered lexically (Rooth, 1992; Kiss, 1998; Carlson, 2014; Lignani; Fonseca, 2023), giving rise to early, strong predictions-by-associations, based on the immediate activation of alternatives (Pickering; Gambi, 2018). In clefts, by contrast, exhaustivity emerges through discourse-level inference (Horn, 1981; Velleman *et al.*, 2012; Pollard; Yasavul, 2016; Onea, 2019), relying on the reconstruction of a question under discussion and resulting in delayed and more flexible predictive-by-production commitments, that take into account the producers’ intentions. These differences in predictive mechanisms naturally explain the distinct timing and strength of processing effects observed across constructions and languages, as well as the role of structural factors in modulating mismatch detection.

This proposal is summarized in Figure 59, which illustrates how exhaustivity is computed and verified across constructions:

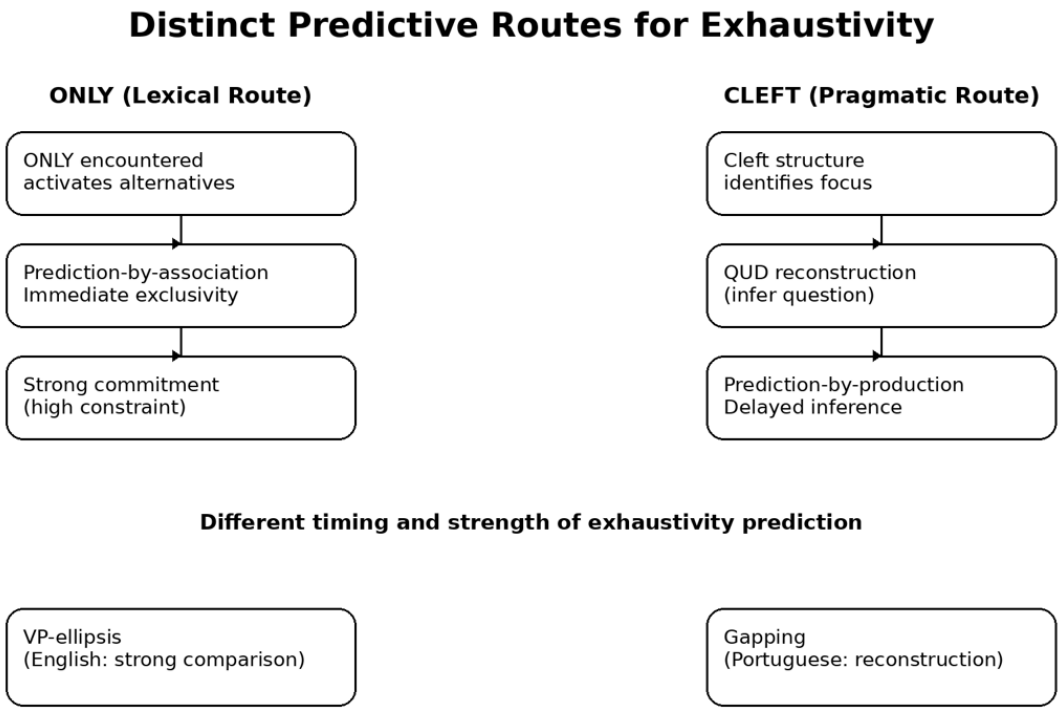


Figure 59 – Distinct predictive routes for exhaustivity processing.

This model aims to integrate the empirical and theoretical observations discussed throughout this thesis. Specifically, it captures the central distinction between two predictive

routes for exhaustivity: a lexically-driven mechanism in ‘only’-sentences, and a discourse-based, inferential mechanism in clefts.

This proposal highlights that these constructions differ not only in whether exhaustivity is encoded, but in how and when predictive commitments are established. In ‘only’-sentences, exhaustivity is triggered immediately upon encountering the focus particle (König, 1991; Kiss, 1998; Lignani; Fonseca, 2023), leading to strong early constraints on interpretation. In clefts, by contrast, exhaustivity emerges more gradually, as a consequence of QUD reconstruction and pragmatic reasoning (Destruel *et al.*, 2015; Onea, 2019). As a result, predictive commitments are weaker and delayed, and processing costs are redistributed to later stages of integration.

Importantly, the model also accounts for the observed eye-tracking patterns by distinguishing between the generation of predictions and their subsequent verification. While predictive commitments may arise early, their consequences become visible in late measures (regression-based), when incoming *input* conflicts with prior expectations. This distinction allows us to reconcile early predictive mechanisms with late processing effects, providing an unified explanation for the patterns observed across experiments and languages.

REFERENCES

ABBUHL, Rebekha; GASS, Susan; MACKEY, Alison. Experimental Research Design. In: PODESVA, Robert J.; SHARMA, Devyani (ed.). Research Methods in Linguistics. Nova Iorque: Cambridge University Press, 2013. p. 116-134.

ALTMANN, Gerry T. M.; KAMIDE, Yuki. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, v. 73, n. 3, p. 247–264, 1999.

ALVES, Michele Calil dos Santos; OLIVEIRA, Talita Moreira de; PEREIRA, Luana Gomes; ALMEIDA, Paloma Bruna de. Processamento de sentenças clivadas de sujeito e objeto denotando hiponímia e contraste no português do Brasil. *ReVEL*, edição especial n. 10, 2015. [www.revel.inf.br]. Available on: http://www.revel.inf.br/files/8ff34075c75611394b73615c9fecc537.pdf. Access on 10 Feb. 2025.

BAAYEN, Harald R.; MILIN, Petar. Analyzing reaction times. *International Journal of Psychological Research*, v. 3, n. 2, p. 12–28, 2010. Available on: https://www.redalyc.org/pdf/2990/299023506003.pdf. Access on 27 May 2026.

BADDELEY, Alan. Working memory. *Memory*, p. 71-111, 2020. Available on: https://www.sciencedirect.com/science/article/pii/S0960982209021332. Access on 27 may 2026.

BATES, Douglas; MÄCHLER, Martin; BOLKER, Ben; WALKER, Steve. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, v. 67, n. 1, p. 1–48, 2015. Available on: https://www.jstatsoft.org/article/view/v067i01. Access on 27 May 2026.

BEAVER, David I.; CLARK, Brady Z. *Sense and Sensitivity: How Focus Determines Meaning*. Oxford: Wiley-Blackwell, 2008.

BECHARA, Evanildo. *Moderna gramática portuguesa*. 39ª ed. Rio de Janeiro: Nova Fronteira, 2019.

BIEZMA, María; RAWLINS, Kyle. Responding to alternative and polar questions. Linguistics and Philosophy, v. 35, n. 5, p. 361-406, 2012.

BIRCH, Stacy; CLIFTON JR, Charles. Focus, accent, and argument structure: Effects on language comprehension. Language and speech, v. 38, n. 4, p. 365-391, 1995.

BRAGA, Maria Luiza; DE MELO BARBOSA, Elisiene. Construções clivadas no português do Brasil sob uma abordagem funcionalista. Matraga-Revista do Programa de Pós-Graduação em Letras da UERJ, v. 16, n. 24, 2009.

BÜRING, Daniel; KRIZ, Manuel. It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). Semantics and Pragmatics, v. 6, p. 6: 1-29, 2013. DOI: 10.3765/sp.6.6 . Access on 27 May 2026.

CAPLAN, David; WATERS, Gloria S. Verbal working memory and sentence comprehension. *Behavioral and brain Sciences*, v. 22, n. 1, p. 77-94, 1999. Available on: https://www.researchgate.net/profile/Maryellen-Macdonald-2/publication/231945760_Fractio

nated_working_memory_Even_in_pebbles_it%27s_still_a_soup_stone/links/596f37294585153d684e1751/Fractionated-working-memory-Even-in-pebbles-its-still-a-soup-stone.pdf. Access on 27 May 2026.

CARLSON, Katy. Clefting, parallelism, and focus in ellipsis sentences. In: FRAZIER, Lyn; GIBSON, Edward (ed.). Explicit and implicit prosody in sentence processing: studies in honor of Janet Dean Fodor. Cham: Springer, 2015. p. 63-83. DOI: https://doi.org/10.1007/978-3-319-12961-7_4

CARLSON, Katy. Predicting contrast in sentences with and without focus marking. Lingua, Morehead, p. 78-91, jul. 2014. Available on: https://www.sciencedirect.com/science/article/pii/S0024384114001624. Access on: 10 Feb. 2026.

CARLSON, Katy. The effects of parallelism and prosody in the processing of gapping structures. Language and Speech, v. 44, n. 1, p. 1-26, 2001. Available on: https://www.researchgate.net/publication/11910942_The_Effects_of_Parallelism_and_Prosody_in_the_Processing_of_Gapping_Structures Access on: 10 Feb. 2026. DOI: 10.1177/00238309010440010101

CINQUE, Guglielmo. Adverbs and functional heads: A cross-linguistic perspective. Nova Iorque: *Oxford University Press*, 1999. 288 p. Available on: https://www.academia.edu/58433772/Adverbs_and_Functional_Heads. Access on 27 May 2026.

CLIFTON, Charles, Jr.; BOCK, Jeannine; RADO, Janina. Effects of the Focus Particle *Only* and Intrinsic Contrast on Comprehension of Reduced Relative Clauses. *Language and Speech*, v. 43, n. 3, p. 309–336, 2000. DOI: https://www.sciencedirect.com/science/chapter/edited-volume/abs/pii/B9780080436425500280?via%3Dihub. Access on 27 de maio de 2026.

CLIFTON, Charles, Jr.; STAUB, Adrian; RAYNER, Keith. Eye movements in reading words and sentences. In: VAN GOMPEL, R. P. G. et al. (Eds.). *Eye Movements: A Window on Mind and Brain*.Oxford: Elsevier, 2007. P. 213–234. Available on: https://people.umass.edu/~cec/CliftonStaubRaynerECEM.pdf. Access on 27 May 2026.

COWLES, H. Wind. Processing information structure: Evidence from comprehension and production. Tese (Doutorado), University of California, San Diego, 2003.

COWLES, H. Wind; KLUENDER, Robert; KUTAS, Marta; POLINSKY, Maria. Violations of information structure: An electrophysiological study of answers to wh-questions. *Brain and Language*, v. 102, n. 3, p. 228–242, 2007. Available on: https://app.amanote.com/v4.5.17/research/note-taking?resourceId=5a_xAnQBKQvf0BhiGlRn Access on 27 May 2026.

CRUSCHINA, Silvio. Syntactic movement and focus-associated implicatures in Hungarian. RGG. Rivista di grammatica generativa, v. 41, p. 1-37, 2019. Available on: https://helda.helsinki.fi/server/api/core/bitstreams/1d333096-d838-4895-adea-d9d94031c4a0/content. Access on 27 May 2026.

CUMMINS, Chris; KATSOS Napoleon (eds), *The Oxford Handbook of Experimental Semantics and Pragmatics*, Oxford Handbooks (2019; online edn, Oxford Academic, 9 May 2019) DOI: https://doi.org/10.1093/oxfordhb/9780198791768.001.0001. Accessed on 24 Mar. 2026.

CHRISTIANSON, Kiel; HOLLINGWORTH, Andrew; HALLIWELL, John F.,; FERREIRA, Fernanda. Thematic roles assigned along the garden path linger. *Cognitive psychology*, v. 42, n. 4, p. 368-407, 2001. Available on: https://ferreiralab.faculty.ucdavis.edu/wp-content/uploads/sites/222/2015/05/Christianson-et-al.-2001_LingeringInterpretationsGP_Cog-Psych.pdf . Accessed on 27 May 2026.

DA SILVA, Andressa Christine Oliveira; FONSECA, Aline Alves. Online Sentence Processing During Reading of Temporally Ambiguous Gapping Sentences in Brazilian Portuguese. Caligrama: Revista de Estudos Românicos, v. 26, n. 2, p. 67-87, 2021.

DE CESARE, L. The focusing function of focusing adverbs. *Linguistik Online*, v. 43, n. 3, 2010.

DELONG, Katherine A.; URBACH, Thomas P.; KUTAS, Marta. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nature neuroscience, v. 8, n. 8, p. 1117-1121, 2005.

DESTRUEL, E.; DEVEAUGH-GEISS, J. P. On the interpretation and processing of French c'est-clefts: Evidence from prosody and eye-tracking. *Glossa*, v. 3, n. 1, p. 1–29, 2018.

DESTRUEL, E.; ONEA, E.; VELLEMAN, D.; MEELEN, M.; ZIMMERMANN, M. A cross-linguistic study of the non-at-issue status of exhaustive inferences. *Semantics & Pragmatics*, v. 8, n. 6, p. 1–52, 2015.

DRENHAUS, H.; ZIMMERMANN, M.; VASISHTH, S. Exhaustiveness effects in clefts are not truth-functional: Violations of exhaustivity in German it-clefts versus *only*-sentences. *Journal of Neurolinguistics*, v. 24, p. 320–337, 2010. DOI: 10.1016/j.jneuroling.2011.10.004

FERREIRA, Fernanda; BAILEY, Karl GD; FERRARO, Vittoria. Good-enough representations in language comprehension. Current directions in psychological science, v. 11, n. 1, p. 11-15, 2002.

FERREIRA, Fernanda; CLIFTON JR, Charles. The independence of syntactic processing. Journal of memory and language, v. 25, n. 3, p. 348-368, 1986.

FERREIRA, Fernanda; LOWDER, Matthew W. Prediction, information structure, and good-enough language processing. In: Psychology of learning and motivation. Academic Press, 2016. p. 217-247.

FILIK, Ruth; PATERSON, Kevin; LIVERSEDGE, Simon. The influence of only and even on online semantic interpretation. Psychonomic bulletin e review, v. 16, n. 4, p. 678-683, 2009. Available on: https://www.researchgate.net/publication/26710847_The_influence_of_only_and_even_on_online_semantic_interpretation Access on: Feb. 2022. DOI: 10.3758/PBR.16.4.678

FILIK, R.; PATERSON, K. G.; LIVERSEDGE, S. P. Processing focus information in question–answer dialogues: An eye movement study. *Journal of Memory and Language*, v. 57, p. 229–245, 2007.

FODOR, Janet Dean. Parsing strategies and constraints on transformations. Linguistic inquiry, v. 9, n. 3, p. 427-473, 1978.

FONSECA, Aline Alves; MAIA, Marcus. Na trilha do processamento da linguagem: o uso de rastreadores oculares na análise de dados linguísticos. Métodos experimentais em psicolinguística, p. 55-74, 2022.

FRAUNDORF, Scott H.; WATSON, Duane G.; BENJAMIN, Aaron S. Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of memory and language*, v. 63, n. 3, p. 367-386, 2010. Available on: https://pmc.ncbi.nlm.nih.gov/articles/PMC2935187/pdf/nihms218220.pdf. Access on 27 May 2026.

FRAZIER, L. Sentence processing: A tutorial review.In: COLTHEART, M. (Ed.). *Attention and Performance XII*. Hillsdale,NJ: Lawrence Erlbaum Associates, 1987. P. 559–586.

FRAZIER, L.; CLIFTON, C. Jr. Construal. Cambridge, MA:MIT Press, 1996.

FRAZIER, L.;RAYNER, K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, v. 14, n. 2, p. 178–210, 1982.

GREIF, Markus; SKOPETEAS, Stavros. Correction by focus: Cleft constructions and the cross-linguistic variation in phonological form. Frontiers in Psychology, v. 12, p. 648478, 2021. DOI: 10.3389/fpsyg.2021.648478. Access on 27 May 2026.

HALVORSEN, P.-K. *The syntax and semantics of cleft constructions*. 1978. Tese (Doutorado em Linguística) – University of Texas at Austin, Austin, 1978.

HERSCH, Jolyn; ANDREWS, Sally. Lexical quality and reading skill: Bottom-up and top-down contributions to sentence processing. Scientific Studies of Reading, v. 16, n. 3, p. 240-262, 2012. Available on: https://www.tandfonline.com/doi/abs/10.1080/10888438.2011.564244. Access on: 10 fev. 2022. DOI: https://doi.org/10.1080/10888438.2011.564244

HORN, Laurence R. Exhaustiveness and the semantics of clefts. 1981. Available on: https://www.researchgate.net/profile/Laurence-Horn/publication/246129460_Exhaustiveness_and_the_Semantics_of_Clefts/links/55be62a508aed621de1210f9/Exhaustiveness-and-the-Semantics-of-Clefts.pdf. Access on 27 jul. 2026.

HUDDLESTON, Rodney D.; PULLUM, Geoffrey K.; BAUER, Laurie. The Cambridge grammar of the English language. Cambridge: Cambridge University Press, 2002.

HÜTTNER, M.;VON HEYDEBRANDT, M.; WESSELS, H. Prosodic marking of focus in German clefts.Journal of Phonetics, v. 32, n. 4, p. 435–460, 2004. DOI:10.1016/S0095-4470(04)00030-0.

ILARI, Rodolfo. Sobre advérbios focalizadores. In.:___. (Org.) Gramática do português falado. vol 2. Campinas: Editora da Unicamp, 2002, 193-212.

JUST, Marcel A.; CARPENTER, Patricia A.; WOOLLEY, Jacqueline D. Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, v. 111, n. 2, p. 228, 1982.

LEVY, Roger. Expectation-based syntactic comprehension. *Cognition*, v. 106, n. 3, p. 1126-1177, 2008.

KING, Jonathan.; JUST, Marcel Adam. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, v. 30, n. 5, p. 580–602, 1991. DOI: 10.1016/0749-596X(91)90027-H.

KISS, Katalin. É. Identificational focus versus information focus. *Language*, v. 74, n. 2, p. 245–273, 1998.

KISS, Katalin É. The English cleft construction as a focus phrase. AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4, p. 217-230, 1999.

KÖNIG, Ekkehard. The meaning of focus particles: a comparative perspective. Nova Iorque: Routledge, 1991. 232 p.

KRIFKA, Manfred. At least some determiners aren't determiners. In: TURNER, K. (ed.). *The semantics/pragmatics interface from different points of view*. Oxford: Elsevier, 1999. p. 257–291. (Current Research in the Semantics/Pragmatics Interface, v. 1).

KRIFKA, Manfred. Basic notions of information structure. In: CARSTON, R.; MORGAN, J. (Eds.). Pragmatics and Semantics. London: Palgrave Macmillan, 2008. p. 13–55. DOI: 10.1057/9780230226153_2.

KUPERBERG, Gina. R.; JAEGER, Florian T.. What do we mean by prediction in language comprehension? Language, Cognition and Neuroscience, v. 31, n. 1, p. 32–59, 2016. DOI: 10.1080/23273798.2015.1102299.

KUTAS, Marta; HILLYARD, Steven. Brain potentials during reading reflect word expectancy and semantic association. Nature, v. 307, n. 5947, p. 161-163, 1984. Available on: https://pubmed.ncbi.nlm.nih.gov/6690995/. Access on: 10 Feb. 2026. DOI: 10.1038/307161a0

KUTAS, Marta; VAN PETTEN, Cyma. K. Psycholinguistics electrified II, 1994–2005. In: KLUENDER, Keith Raymond. Handbook of psycholinguistics. Academic Press, 2006. p. 83-143. Available on: http://kutaslab.ucsd.edu/people/kutas/pdfs/1994.HP.83.pdf. Access on: Feb. 2026.

LAMBRECHT, Knud. Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge: *Cambridge University Press*, 1994.

LEVY, Roger. Memory and surprisal in human sentence comprehension. In: VAN GOMPEL, Roger (ed.). Sentence processing. Londres: Psychology Press, 2013. p. 78-114. Available on: https://psycnet.apa.org/record/2013-32124-004. Access on: 10 Feb. 2022.

LIGNANI, Bianca Chaves Leite; FONSECA, Aline Alves. Exaustividade com advérbios focalizadores e clivadas totais em Português Brasileiro. *Veredas-Revista de Estudos Linguísticos*, v. 27, n. 1, 2023. DOI: https://doi.org/10.34019/1982-2243.2023.v27.40747

LIKERT, Rensis. A technique for the measurement of attitudes. Archives of psychology, 1932.

LIVERSEDGE, Simon; PATERSON, Kevin; CLAYES, Emma. The influence of only on syntactic processing of "long" relative clause sentences. The Quarterly Journal of Experimental Psychology Section A, v. 55, n. 1, p. 225-240, 2002. Available on: https://journals.sagepub.com/doi/10.1080/02724980143000253. Access on: 10 Feb. 2022.

MAIA, Marcus. Processos bottom-up e top-down no rastreamento ocular de imagens. Veredas-Revista de Estudos Linguísticos, p. 8-24, 2008.

MARTIN, Andrea E.; MCELREE, Brian. Memory operations that support language comprehension: evidence from verb-phrase ellipsis. Journal of Experimental Psychology: Learning, Memory, and Cognition, v. 35, n. 5, p. 1231, 2009.

NIEUWLAND, M. S.; VAN BERKUM, J. J. A. When peanuts fall in love: N400 evidence for the power of discourse. Journal of Cognitive Neuroscience, v. 18, n. 7, p. 1098–1111, 2006. DOI: 10.1162/jocn.2006.18.7.1098.

ONEA, Edgar. Exhaustivity in It -Clefts. T*he Oxford Handbook of Experimental Semantics and Pragmatics*, v. 1, 401-417. 2019. University Press. Oxford DOI: https://doi.org/10.1093/oxfordhb/9780198791768.013.17

PATERSON, Kevin; LIVERSEDGE, Simon; FILIK, Ruth; JUHASZ, Barbara; WHITE, Sarah; RAYNER, Keith. Focus identification during sentence comprehension: Evidence from eye movements. Quarterly Journal of Experimental Psychology, v. 60, n. 10, p. 1423-1445, 2007. Available on: https://www.tandfonline.com/doi/abs/10.1080/17470210601100563. Access on: 10 Feb. 2022.

PATERSON, Kevin; LIVERSEDGE, Simon; UNDERWOOD, Geoffrey. The influence of focus operators on syntactic processing of short relative clause sentences. The Quarterly Journal of Experimental Psychology Section A, v. 52, n. 3, p. 717-737, 1999. Available on: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.332.4696&rep=rep1&type=pdf. Access on: 10 Feb. 2022.

PICKERING, Martin; GAMBI, Chiara. Predicting while comprehending language: A theory and review. Psychological Bulletin, v. 144, n. 10, p. 1002, 2018. Available on: https://pubmed.ncbi.nlm.nih.gov/29952584/. Access on: 10 Feb. 2022. DOI: 10.1037/bul0000158.

PLIATSIKAS, Christos; MARINIS, Theodoros. Online psycholinguistic methods in Second language acquisition research. *The Encyclopedia of Applied Linguistics*, p. 1-9, 2012.

POLLARD, Carl; YASAVUL, Murat. Anaphoric it-clefts: The myth of exhaustivity. Proceedings of CLS, 2016.

QUAREZEMIN, Sandra; TESCARI NETO, A. Da sintatização dos focos contrastivo e exaustivo em CP e das estratégias de marcação de foco. *ReVEL*, Mercado de Letras, edição especial, n. 10, p. 42-77, 2015.

RAYNER, Keith. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, v. 124, n. 3, p. 372–422, 1998.

REIS, S. A interface entre estrutura informacional e sintaxe: um estudo comparativo entre o português europeu e o português brasileiro. *Revista da ABRALIN*, v. 12, n. 2, p. 291–318, 2013.

R Core Team (2026). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

ROOTH, Mats. A theory of focus interpretation. Natural language semantics, v. 1, n. 1, p. 75-116, 1992. Available on: https://link.springer.com/article/10.1007/BF02342617 Access on: 10 Feb. 2026.

RYSKIN, Rachel; LEVY, Roger; FEDORENKO, Evelina. Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. Neuropsychologia, v. 136, p. 107258, 2020. Available on: https://pubmed.ncbi.nlm.nih.gov/31730774/ . Access on: 10 Feb. 2026. DOI: 10.1016/j.neuropsychologia.2019.107258

SANTOS, H. R. Foco e estrutura informacional: clivadas em português. *DELTA*, v. 18, n. 2, p. 285–309, 2002.

SKOPETEAS, Stavros.; FANSELOW, Gisbert. Alternatives in focus and givenness: Experimental evidence. Journal of Semantics, v. 27, n. 1, p. 1–40, 2010. DOI: 10.1093/jos/ffp015.

SILVEIRA, Damaris Matias. Sentenças clivadas no português brasileiro: aspectos formais de estruturas focalizadoras à luz do programa cartográfico. *Cadernos de Estudos Linguísticos*, v. 63, p. e021023-e021023, 2021.

SOUZA, Edson Rosa Francisco de. Os advérbios focalizadores no português falado do Brasil: uma abordagem funcionalista. Dissertação (Mestrado) – UNESP, São Paulo, 2004.

STAUB, Adrian. The effect of lexical predictability on eye movements in reading. Journal of Experimental Psychology: Human Perception and Performance, v. 41, n. 5, p. 1285–1304, 2015. DOI: 10.1037/xhp0000109.

TRAXLER, Matthew. Sentence Processing. In: Traxler, M.J. Introduction to psycholinguistics: Understanding language science. 2013.

VAN DYKE, J. A.; MCELREE, B. Retrieval interference in sentence comprehension. Journal of Memory and Language, v. 55, n. 2, p. 157–166, 2006. DOI: 10.1016/j.jml.2006.03.007.

VAN GOMPEL, R. P. G.; MAJID, A. Focus and the syntactic structure of clefts. Linguistics, v. 42, n. 6, p. 1119–1145, 2004. DOI: 10.1515/ling.2004.42.6.1119.

VAN HERTEN, M.; KOLK, H. H. J.; CHWILLA, D. J. When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, v. 18, n. 7, p. 1181–1197, 2006.

VELLEMAN, Dan; BEAVER, David; DESTRUEL, Emilie; BUMFORD, Dylan; ONEA, Edgar; COPPICK, Liz. It-clefts are IT (inquiry terminals). In: Proceedings of SALT 22, p. 441–460, 2012. DOI: 10.3765/salt.v22i0.2507.

WILSON, Robert A.; KEIL, Frank C. (Ed.). *The MIT Encyclopedia of the cognitive sciences* (MITECS). MIT press, 2001.

ZEHR, Jeremy; SCHWARZ, Florian. Penncontroller for internet based experiments (ibex). URL https://doi. org/10, v. 17605, 2018.

ZIMMERMAN, Malte. The processing of focus alternatives in German and English clefts. Lingua, v. 114, n. 8, p. 901–921, 2004. DOI: 10.1016/S0024-3841(03)00063-0.

ZIMMERMANN, Malte; VEAUGH-GEISS, Joseph P de..; TÖNNIS, Swantje; ONEA, Edgar. (Non-) exhaustivity in focus partitioning across languages. 2020.

APPENDIX A – EXPERIMENTAL ITEMS OF AJT1

1. At Christmas (only/it was) the grandmother (who) prepared the meals and (the grandfather did too/the drinks too).
2. At weddings (only/it was) the musician (who) had a break and (the photographer did too/a snack too).
3. Every morning (only/it was) Alice (who) made the coffee and (Mary did too/the tea too).
4. On Tuesdays (only/it was) the manager (who) closed the store and (the salesman did too/the warehouse too).
5. At the clinic (only/it was) the doctor (who) prescribes the pills and (the nurse did too/the ointments too).
6. For the party (only/it was) Gabriel (who) brought some wine and (Brandon did too/a cake too).
7. During the press conference (only/it was) the journalist (who) used a notebook and (the policeman did too/a microphone too).
8. On Wednesday (only/it was) the accountant (who) signed the request and (the surveyor did too/the certificate too).
9. Last week (only/it was) the maid (who) cleaned the bedroom and (the cleaning lady did too/the bathroom too).
10. Before the football match (only/it was) the goalkeeper (who) stretched his legs and (the referee did too/his arms too).
11. At the formal dinner (only/it was) the judge (who) tried the fish and (the prosecutor did too/the soup too).
12. After the investigation (only/it was) the manager (who) seemed worried and (the accountant did too/regretful too).
13. At the rehearsal (only/it was) the actor (who) recited a poem and (the writer did too/a limerick too).
14. During the tea (only/it was) the grandmother (who) ate a sandwich and (the mother did too/some cake too).
15. During the summer (only/it was) Phillip (who) tidied the garden and (the neighbor did too/the lawn too).
16. At the department store (only/it was) Lucy (who) bought some fabric and (Margaret did too/some paint too).
17. At the casino (only/it was) William (who) played poker and (Robert did too/blackjack too).
18. After the lecture (only/it was) Beatrice (who) read some articles and (Samantha did too/a book too).
19. In the orchestra (only/it was) Nicholas (who) played the trombone and (Samuel did too/the trumpet too).
20. After dinner (only/it was) Andrew (who) cleaned the dishes and (Karen did too/the cutlery too).
21. In the circus (only/it was) the clown (who) dyed his hair and (the acrobat did too/his eyebrows too).
22. At the library (only/it was) Lucy (who) read the books and (Claire did too/comics too).
23. At the mall (only/it was) Layla (who) bought clothes and (Natalie did too/shoes too).

24. At the cinema (only/it was) Felicity (who) drank soda and (Meredith did too/water too).
25. In the summer (only/it was) Patricia (who) drove to the beach and (Lorraine did too/to the lake too).
26. During the Olympics (only/it was) Edith (who) watched volleyball and (Wendy did too/basketball too).
27. In the album (only/it was) the singer (who) wrote the lyrics and (the guitarist did too/the chords too).
28. At the park (only/it was) the child (who) petted a dog and (his mother did too/a cat too).
29. At the café (only/it was) Olivia (who) ordered a coffee and (Kate did too/a muffin too).
30. At the restaurant (only/it is) the chef (who) selects the ingredients and (the owner did too/the wines too).
31. During college (only/it was) William (who) studied electives and (Matt did too/music too).
32. For the picnic (only/it was) Polly (who) brought juice and (Lauren did too/fruits too).
33. At the show (only/it was) Oliver (who) enjoyed the clowns and (Jennifer did too/the jugglers too).
34. For the ball (only/it was) Samantha (who) dressed in a red skirt and (Mindy did too/a red top too).
35. During the business meeting (only/it was) Michael (who) discussed finances and (Christine did too/office space too).
36. At the concert (only/it was) the violinist (who) played the solo and (the drummer did too/the interlude too).
37. In the last chapter (only/it was) the protagonist (who) drank the poison and (the villain did too/the wine too).
38. While preparing breakfast (only/it was) Mary (who) fried the eggs and (Gareth did too/the bacon too).
39. At the bar (only/it was) Robert (who) drank some beer and (Barbara did too/some whisky too).
40. After the party (only/it was) Zack (who) found his jacket and (Helen did too/his keys too).
41. After the interview (only/it was) Daniel (who) called security and (Christian did too/the police too).
42. During the fight (only/it was) the mother (who) broke the glass and (the father did too/the cup too).
43. In the movie (only/it was) the actress (who) wore make-up and (the actor did too/a wig too).
44. During the competition (only/it was) the Argentinian (who) passed out and (the English did too/vomited too).
45. Every Friday (only/it is) the Math teacher (who) plans the lessons and (the English teacher did too/the exercises too).
46. Last week (only/it was) Dave (who) painted an egg and (Lori did too/a ball too).
47. This week (only/it was) Kenny (who) wiped the table and (his mother did too/the floor too).
48. Last night (only/it was) Eric (who) caught a train and (Paul did too/a bus too).

APPENDIX B – EXPERIMENTAL ITEMS OF AJT2

1. At Christmas (only/it was) the grandmother (who) prepared the meals and also (the grandfather/the drinks).
2. At weddings (only/it was) the musician (who) had a break and also (the photographer/a snack).
3. Every morning (only/it was) Alice (who) made the coffee and also (Mary/the tea).
4. On Tuesday (only/it was) the manager (who) closed the store and also (the salesman/the warehouse).
5. At the clinic (only/it was) the doctor (who) prescribed the pills and also (the nurse/the ointments).
6. For the party (only/it was) Gabriel (who) brought some wine and also (Brandon/a cake).
7. During the press conference (only/it was) the journalist (who) used a notebook and also (the policeman/a microphone).
8. On Wednesday (only/it was) the accountant (who) signed the request and also (the surveyor/the certificate).
9. Last week (only/it was) the maid (who) cleaned the bedroom and also (the cleaning lady/the bathroom).
10. Before the football match (only/it was) the goalkeeper (who) stretched his legs and also (the referee/his arms).
11. At the formal dinner (only/it was) the judge (who) ordered the fish and also (the prosecutor/the soup).
12. After the investigation (only/it was) the manager (who) seemed worried and also (the accountant/regretful).
13. At the rehearsal (only/it was) the actor (who) recited a poem and also (the writer/a limerick).
14. During the tea (only/it was) the grandmother (who) ate a sandwich and also (the mother/some cake).
15. During the summer (only/it was) Phillip (who) tidied the garden and also (his neighbour/the lawn).
16. At the department store (only/it was) Lucy (who) bought some fabric and also (Margaret/some paint).
17. At the casino (only/it was) William (who) played poker and also (Robert/blackjack).
18. After the lecture (only/it was) Beatrice (who) read some articles and also (Samantha/a book).
19. In the orchestra (only/it was) Nicholas (who) played the trombone and also (Samuel/the trumpet).
20. After dinner (only/it was) Andrew (who) cleaned the dishes and also (Karen/the cutlery).
21. In the circus (only/it was) the clown (who) dyed his hair and also (the acrobat/his eyebrows).
22. At the library (only/it was) Lucy (who) read the books and also (Claire/comics).
23. At the mall (only/it was) Layla (who) bought clothes and also (Natalie/shoes).
24. At the cinema (only/it was) Felicity (who) drank soda and also (Meredith/water).
25. In the summer (only/it was) Patricia (who) drove to the beach and also (Lorraine/to the lake).

26. During the Olympics (only/it was) Edith (who) watched volleyball and also (Wendy/basketball).
27. In the album (only/it was) the singer (who) wrote the lyrics and also (the guitarist/the chords).
28. At the park (only/it was) the child (who) petted a dog and also (his mother/a cat).
29. At the café (only/it was) Olivia (who) ordered a coffee and also (Kate/a muffin).
30. At the restaurant (only/it is) the chef (who) selects the ingredients and also (the owner/the wines).
31. During college (only/it was) William (who) studied electives and also (Matt/music).
32. For the picnic (only/it was) Polly (who) brought juice and also (Lauren/fruits).
33. At the show (only/it was) Oliver (who) enjoyed the clowns and also (Jennifer/the jugglers).
34. For the ball (only/it was) Samantha (who) dressed in a red skirt and also (Mindy/a red top).
35. During the business meeting (only/it was) Michael (who) discussed finances and also (Christine/office space).
36. At the concert (only/it was) the violinist (who) played the solo and also (the drummer/the interlude).
37. In the last chapter (only/it was) the protagonist (who) drank the poison and also (the villain/the wine).
38. While preparing breakfast (only/it was) Mary (who) fried the eggs and also (Gareth/the bacon).
39. At the bar (only/it was) Robert (who) drank some beer and also (Barbara/some whisky).
40. After the party (only/it was) Zack (who) found his jacket and also (Helen/his keys).
41. After the interview (only/it was) Daniel (who) called security and also (Christian/the police).
42. During the fight (only/it was) the mother (who) broke the glass and also (the father/the cup).
43. In the movie (only/it is) the actress (who) wears make-up and also (the actor/a wig).
44. During the competition (only/it was) the Argentinian (who) passed out and also (the English/vomited).
45. Every Friday (only/it is) the Math teacher (who) plans the lessons and also (the English teacher/the exercises).
46. Last week (only/it was) Dave (who) painted an egg and also (Lori/a ball).
47. This week (only/it was) Kenny (who) wiped the table and also (his mother/the floor).
48. Last night (only/it was) Eric (who) caught a train and also (Paul/a bus).

APPENDIX C – EXPERIMENTAL ITEMS OF AJT3

1. No feriado (só/foi) o vigia (quem) trabalhou na universidade e também (o faxineiro/em casa).
2. Em casamentos (só/foi) o músico (quem) atrasou e também (o fotógrafo/bebeu).
3. Diariamente (só/foi) a Andressa (quem) fez café e também (a Mariana/chá).
4. Na terça-feira (só/foi) o gerente (quem) fechou a loja e também (o vendedor/o estoque).
5. Esse ano (só/foi) o clínico (quem) prescreveu o comprimido e também (o cirurgião/a pomada).
6. Para a festa (só/foi) o Gabriel (quem) comprou um presente e também (o Leonardo/um bolo).
7. Durante a coletiva (só/foi) o jornalista (quem) usou o gravador e também (o jogador/o microfone).
8. Na quarta-feira (só/foi) o noivo (quem) assinou o pedido e também (o padre/a certidão).
9. Inicialmente (só/foi) a camareira (quem) limpou o quarto e também (a faxineira/o banheiro).
10. Durante o jogo (só/foi) a goleira (quem) alongou as pernas e também (a zagueira/as costas).
11. No jantar (só/foi) a juíza (quem) experimentou o peixe e também (a promotora/o purê).
12. Em dezembro (só/foi) a diretora (quem) suspendeu Diogo e também (a coordenadora/Henrique).
13. No sábado (só/foi) o policial (quem) escreveu o relatório e também (o bombeiro/a ocorrência).
14. No almoço de Páscoa (só/foi) a avó (quem) serviu o camarão e também (a mãe/o bacalhau).
15. Depois da tempestade (só/foi) Felipe (quem) aparou o jardim e também (o vizinho/o gramado).
16. Ontem (só/foi) a Letícia (quem) recortou o papel e também (a Marcela/o tecido).
17. Nas férias (só/foi) o Otávio (quem) jogou baralho e também (o Rodrigo/Uno).
18. Durante a aula (só/foi) a Beatriz (quem) terminou os exercícios e também (a Renata/a leitura).
19. Supostamente (só/foi) o Nícolas (quem) tocou baixo e também (o Samuel/bateria).
20. Todo final de semana (só/foi) o Tiago (quem) lavou a louça e também (a Carina/o chão).
21. Em qualquer voo (só/foi) a Patrícia (quem) usou o cinto e também (a Marina/o fone de ouvido).
22. Na escola (só/foi) a Lavínia (quem) guardou os brinquedos e também (a Clara/os livros).
23. No shopping (só/foi) a Jéssica (quem) comprou roupas e também (a Nicole/sapatos).
24. No cinema (só/foi) a Fernanda (quem) bebeu refrigerante e também (a Mirela/água).

APPENDIX D – EXPERIMENTAL ITEMS OF THE EYE-TRACKING TASK IN ENGLISH

	Context sentences	Target sentences
1.	Anna and Olivia run a busy bakery in London.	Yesterday (only/it was) Anna (who) brewed coffee and (Olivia did/pastries) too.
2.	Sam and Jake teach at a school in Nottingham.	Last week (only/it was) Sam (who) assigned homework and (Jake did/projects) too.
3.	Lisa and Maria explore different cities around the world.	Last month (only/it was) Lisa (who) visited museums and (Maria did/parks) too.
4.	Kevin and Tom host a popular podcast about travelling.	Last episode (only/it was) Kevin (who) interviewed a guide and (Tom did/a taxi driver) too.
5.	Emma and Chloe design stunning websites for clients.	This morning (only/it was) Emma (who) created a layout and (Chloe did/graphics) too.
6.	Noah and Clint manage a bustling farm in the countryside.	Last week (only/it was) Noah (who) harvested vegetables and (Clint did/fruits) too.
7.	Zoe and Cindy run a boutique clothing store in their town.	Yesterday (only/it was) Zoe (who) organised new arrivals and (Cindy did/the shelves) too.
8.	Daniel and Bill produce an indie film in their hometown.	Last weekend (only/it was) Daniel (who) filmed scenes and (Bill did/landscapes) too.
9.	Jack and Lyn run a cozy bookstore in the city.	This morning (only/it was) Jack (who) organised the shelves and (Lyn did/the displays) too.
10.	Paul and Oliver lead a community yoga class in the park.	Last week (only/it was) Paul (who) demonstrated poses and (Oliver did/breathing exercises) too.

11.	Lily and Daisy create beautiful art for local galleries.	Yesterday (only/it was) Lily (who) painted a mural and (Daisy did/crafted sculptures) too.
12.	Jake and Stephen volunteer at an animal shelter in their neighbourhood.	Last Saturday (only/it was) Jake (who) fed the dogs and (Stephen did/the cats) too.
13.	Grace and Ava run a tech startup focused on innovation.	This week (only/it was) Grace (who) developed some software and (Ava did/an app) too.
14.	Ryan and Leo coach a youth soccer team in their town.	Last practice (only/it was) Ryan (who) taught strategies and (Leo did/techniques) too.
15.	Hannah and Vicky write a blog about sustainable living.	Recently (only/it was) Hannah (who) researched eco-friendly products and (Vicky did/some sustainability tips) too.
16.	Ed and Max work at a fitness studio in their city.	This week (only/it was) Ed (who) led a spin class and (Max did/a yoga class) too.
17.	Evelyn and Lydia own a flower shop in the neighbourhood.	Yesterday (only/it was) Evelyn (who) arranged bouquets and (Lydia did/centerpieces) too.
18.	Liam and Aiden organise community events in their town.	Last month (only/it was) Liam (who) planned a festival and (Aiden did/a talent show) too.
19.	Tyler and Carl create mobile apps for small businesses.	Recently (only/it was) Tyler (who) tested performance and (Carl did/functionality) too.
20.	Tiffany and Mia have a popular YouTube channel about cooking.	Last week (only/it was) Tiffany (who) uploaded a recipe and (Mia did/a food review) too.

21.	Ella and Regina host a podcast about mental health awareness.	Last episode (only/it was) Ella (who) shared some tips and (Regina did/some stories) too.
22.	Sophia and Cynthia teach dance classes at a local studio.	Last night (only/it was) Sophia (who) choreographed a routine and (Cynthia did/a presentation) too.
23.	Ethan and Chris run a podcast about technology trends.	Last week (only/it was) Ethan (who) presented innovations and (Chris did/new startups) too.
24.	Ben and Caleb run a photography business specialising in portraits.	Yesterday (only/it was) Ben (who) shot a family session and (Caleb did/an editorial) too.
25.	Joan and Betty organise art workshops for children.	Last weekend (only/it was) Joan (who) prepared materials and (Betty did/activities) too.
26.	Peggy and Hannah work at a travel agency in the city.	One hour ago (only/it was) Peggy (who) booked flights and (Hannah did/accommodations) too.
27.	Peter and Alex coach a local football team.	Last practice (only/it was) Peter (who) conducted a tactical training and (Alex did/a conditioning session) too.
28.	Marianne and Karen host a blog about outdoor adventures.	Last vacation (only/it was) Marianne (who) hiked a new trail and (Karen did/a mountain) too.
29.	Emily and Gabrielle run a cafè known for its pastries and desserts.	Yesterday (only/it was) Emily (who) baked croissants and (Gabrielle did/cakes) too.
30.	Landon and Carl organise charity events for local causes.	Last event (only/it was) Landon (who) coordinated activities and (Carl did/donations) too.
31.	Orion and Dorian run a mobile car wash service.	Last week (only/it was) Orion (who) cleaned vehicles and (Dorian did/interiors) too.

32.	Zane and Adrian have a dog training business.	Last session (only/it was) Zane (who) taught obedience commands and (Adrian did/socialisation techniques) too.
33.	Allison and Ivy create educational videos for kids.	Last week (only/it was) Allison (who) wrote the scripts and (Ivy did/the lesson plans) too.
34.	Layla and Myriam lead a book club in their community.	Last month (only/it was) Layla (who) selected the book and (Myriam did/the place of the meeting) too.
35.	Axel and Kyle run a photography studio focused on events.	Last weekend (only/it was) Axel (who) covered a wedding and (Kyle did/a family reunion) too.
36.	Liora and Nia organise cultural festivals in their town.	Last festival (only/it was) Liora (who) arranged vendors and (Nia did/locations) too.
37.	Rowan and Jareth have a YouTube channel focused on DIY projects.	Last video (only/it was) Rowan (who) filmed a tutorial and (Jareth did/the artifact) too.
38.	Astrid and Lyra run a tutoring service for students.	Last week (only/it was) Astrid (who) created lesson plans and (Lyra did/a workbook) too.
39.	Marina and Selene run a sustainable fashion brand.	Last year (only/it was) Marina (who) designed a new collection and (Selene did/a new campaign) too.
40.	Jason and Finley host a podcast about personal finance.	Last book (only/it was) Jason (who) researched topics and (Finley did/the summary) too.
41.	Veda and Thalia produce educational content for children.	Last month (only/it was) Veda (who) created stories and (Thalia did/illustrations) too.

42.	Justin and Leif run a tech support business for seniors.	Last week (only/it was) Justin (who) helped clients and (Leif did/members of the staff) too.
43.	Iris and Luna create a newsletter about local events.	Two weeks ago (only/it was) Iris (who) designed a logo and (Luna did/the layout) too.
44.	Kai and Bodhi run a graphic design agency.	Last week (only/it was) Kai (who) worked on logos and (Bodhi did/on branding) too.
45.	Fiona and Elena organise hiking trips for enthusiasts.	Last month (only/it was) Fiona (who) planned routes and (Elena did/a hiking guide) too.
46.	Jasper and Magnus manage a YouTube channel focused on fitness.	Last month (only/it was) Jasper (who) filmed workouts and (Magnus did/dance classes) too.
47.	Sophie and Maris run a co-working space in their city.	Last year (only/it was) Sophie (who) managed events and (Maris did/memberships) too.
48.	Lena and Jasmin run a home cleaning service.	Last weekend (only/it was) Lena (who) cleaned houses and (Jasmin did/apartments) too.
49.	Lyra and Ophelia run a small event planning business.	Last week (only/it was) Lyra (who) organised venues and (Ophelia did/guest lists) too.
50.	Maximus and Connell organise sports leagues for youth.	Last season (only/it was) Maximus (who) recruited players and (Connell did/recruited coaches) too.
51.	Teresa and Nova run an art studio for aspiring artists.	Last week (only/it was) Teresa (who) hosted workshops and (Nova did/exhibitions) too.

1.

52.	Samuel and Elijah run a tech consultancy for startups.	Yesterday (only/it was) Samuel (who) analysed client needs and (Elijah did) too.
53.	Naomi and Isla create an online course for language learners.	Last week (only/it was) Naomi (who) designed a book and (Isla did/quizzes) too.
54.	Isaac and Felix manage a local restaurant.	Last weekend (only/it was) Isaac (who) prepared the mains and (Felix did/the desserts) too.
55.	Monica and Izzie operate a mobile coffee truck.	This week (only/it was) Monica (who) served coffee and (Izzie did/cappuccino) too.
56.	Owen and Richard develop a lifestyle blog about minimalism.	Last week (only/it was) Owen (who) wrote articles and (Richard did/comments) too.
57.	Maya and Elara organise charity runs for health awareness.	Last year (only/it was) Maya (who) planned routes and (Elara did/events) too.
58.	Miles and Jude create a mobile app for budgeting.	Last month (only/it was) Miles (who) developed the user interface and (Jude did/the features) too.
59.	Lila and Daphne run a garden centre.	Last week (only/it was) Lila (who) selected plants and (Daphne did/displays) too.
60.	Cassian and Eamon manage a digital marketing agency.	Last month (only/it was) Cassian (who) developed campaigns and (Eamon did/brandings) too.
61.	Isabella and Vivienne run a podcast about environmental issues.	Last week (only/it was) Isabella (who) interviewed activists and (Vivienne did/politicians) too.
62.	Hunter and Chase operate a local delivery service.	Last month (only/it was) Hunter (who) coordinated routes and (Chase did/drivers) too.

63.	Martina and Julia create a buffet experience for events.	Last month (only/it was) Martina (who) curated the menu and (Julia did/the drink menu) too.
64.	Marco and Luca coach a youth cricket team.	Last week (only/it was) Marco (who) led training drills and (Luca did/the practice) too.
65.	William and Henry own a music production studio.	Last month (only/it was) William (who) produced tracks and (Henry did/beats) too.
66.	Amina and Samira organise yoga retreats.	Last year (only/it was) Amina (who) organised classes and (Samira did/organised accommodation) too.
67.	Ravi and Ian work at a healthcare clinic.	Last weekend (only/it was) Ravi (who) administered medicines and (Ian did/treatments) too.
68.	Alice and Amelia are surgeons at a hospital.	Yesterday (only/it was) Alice (who) performed surgeries and (Amelia did/post-operative assessments) too.
69.	James and Charles run a fitness studio.	Last month (only/it was) James (who) offered workout routines and (Charles did/nutritional advice) too.
70.	Kate and Charlotte organise community clean-up events.	Recently (only/it was) Kate (who) coordinated volunteers and (Charlotte did/activities) too.
71.	Thomas and George create a web series about technology.	Last month (only/it was) Thomas (who) edited some scripts and (George did/some footage) too.
72.	Elizabeth and Eleanor run a language school.	Last year (only/it was) Elizabeth (who) developed a pedagogical method and (Eleanor did/curricula) too.
73.	Dexter and Quinn create video games.	Last week (only/it was) Dexter (who) developed characters and (Quinn did/the storylines) too.

74.	Mary and Lucy lead workshops on public speaking.	Last week (only/it was) Mary (who) prepared materials and (Lucy did/lessons) too.
75.	Soren and Zephyr operate a farm-to-table restaurant.	Last weekend (only/it was) Soren (who) bought ingredients and (Zephyr did/kitchenware) too.

1.

76.	Claire and Maggie own an online clothing store.	Last month (only/it was) Claire (who) managed sponsors and (Maggie did/inventory) too.
77.	Gabriel and Jackson have a podcast about storytelling.	Last episode (only/it was) Gabriel (who) read personal stories and (Jackson did/literature classics) too.
78.	Arya and Rachel own a consulting business for non-profits.	Last month (only/it was) Arya (who) assessed needs and (Rachel did/goals) too.
79.	Nicholas and Hugh produce a music festival.	Last year (only/it was) Nicholas (who) coordinated activities and (Hugh did/logistics) too.
80.	Melissa and Mindy have a lifestyle Instagram profile.	Last weekend (only/it was) Melissa (who) posted stories and (Mindy did/reels) too.
81.	Michael and Phillip develop a coding bootcamp.	Last month (only/it was) Michael (who) planned learning outcomes and (Phillip did/assignments) too.
82.	Demi and Rose organise film screenings for the community.	Last week (only/it was) Demi (who) selected films and (Rose did/places) too.
83.	Andy and Jim write an online magazine about Literature.	Last month (only/it was) Andy (who) wrote articles and (Jim did/reviews) too.

84.	Angela and Jill host a podcast about entrepreneurship.	Last week (only/it was) Angela (who) interviewed founders and (Jill did/clients) too.
85.	Joe and Randolph like giving parties.	Last party (only/it was) Joe (who) invited friends and (Randolph did/coworkers) too.
86.	Polly and Cathy have a YouTube channel about music and bands.	Last month (only/it was) Polly (who) researched music genres and (Cathy did/new bands) too.
87.	Jerry and Duncan produce a documentary series.	Last year (only/it was) Jerry (who) managed shootings and (Duncan did/personnel) too.
88.	Vera and Samantha run an organic farming business.	Last year (only/it was) Vera (who) planned crop rotations and (Samantha did/sustainable farming practices) too.
89.	Mika and Alek manage a pet adoption centre.	Last week (only/it was) Mika (who) organised adoption events and (Alek did/dog food packages) too.
90.	Nathan and Walter have a ship chartering company.	Last month (only/it was) Nathan (who) arranged shipping routes and (Walter did/crew schedules) too.
91.	Nathalie and Valerie own a wedding planning business.	Last wedding (only/it was) Nathalie (who) decided the menu and (Valerie did/the theme) too.
92.	Elliot and Daniel run an independent record label.	Last year (only/it was) Elliot (who) signed new singers and (Daniel did/new sponsors) too.
93.	Jean and Lara work in a home renovation company.	Last week (only/it was) Jean (who) met with clients and (Lara did/masons) too.
94.	Arthur and Frederick have a gourmet food truck.	Last night (only/it was) Arthur (who) managed food delivery apps and (Frederick did/customer orders) too.

95.	Beatrice and Penelope organise the school talent show.	Last weekend (only/it was) Beatrice (who) came up with performance ideas and (Penelope did/posters) too.
96.	Arlo and Victor have an online course for aspiring chefs.	Last month (only/it was) Arlo (who) filmed recipes and (Victor did/cooking demonstrations) too.
97.	Vivian and Allie produce a series of workshops for writers.	Last month (only/it was) Vivian (who) planned topics and (Allie did/materials) too.
98.	Stacey and Jane coordinate an animal rescue and rehabilitation centre.	Last week (only/it was) Stacey (who) cared for healthy animals and (Jane did/injured animals) too.
99.	Leonard and Florian run a classic car restoration business.	Last month (only/it was) Leonard (who) sourced rare car parts and (Florian did/wheels) too.
100.	Barbara and Nina run an architectural design firm.	Last week (only/it was) Barbara (who) designed floor plans and (Nina did/furniture) too.
101.	Alistair and Bernard are working on a history project together.	Last week (only/it was) Alistair (who) researched World War I and (Bernard did/World War II) too.
102.	Oscar and Bill work at a vintage watch repair shop.	Last month (only/it was) Oscar (who) repaired watches and (Nill did/clocks) too.
103.	Phoebe and Patricia play in a band together.	Last year (only/it was) Phoebe (who) wrote lyrics and (Patricia did/chords) too.
104.	Stuart and Rick own a traditional pub.	Yesterday (only/it was) Stuart (who) served the drinks and (Rick did/the meals) too.
105.	Megan and Ellis host a music podcast.	Last episode (only/it was) Megan (who) suggested artists and (Ellis did/new albums) too.

106.	Dylan and Rhys are lawyers in a law firm.	Last month (only/it was) Dylan (who) drafted legal contracts and (Rhys did/patent applications) too.
107.	Carys and Eira developed a language learning app.	Last month (only/it was) Carys (who) improved lessons and (Eira did/fixation games) too.
108.	Edith and Caitlin manage a boutique clothing store.	Last weekend (only/it was) Edith (who) assisted the salespeople and (Caitlin did/the customers) too.
109.	Gregor and Tony write for a legal advice website.	Yesterday (only/it was) Gregor (who) researched legal topics and (Tony did/tribunal decisions) too.
110.	Sally and Alissa organise community theatre productions.	Last presentation (only/it was) Sally (who) directed rehearsals and (Alissa did/performances) too.
111.	John and Andrew have an online artisanal food shop.	Last month (only/it was) John (who) made pasta and (Andrew did/desserts) too.
112.	Denise and Judith run an eco-friendly laundry service.	Yesterday (only/it was) Denise (who) checked the driers and (Judith did/the washing cycles) too.
113.	Matthew and Keith work at a hospital.	Last month (only/it was) Matthew (who) coordinated patient care and (Keith did/surgeries) too.
114.	Steve and Donald work on an engineering project.	Last project (only/it was) Steve (who) developed prototypes and (Donald did/tests) too.
115.	Cecilia and Trudy run a tree care and landscaping business.	Last week (only/it was) Cecilia (who) supervised tree trimming and (Trudy did/planting) too.
116.	Jackie and Mona own a social media influencer agency.	Last month (only/it was) Jackie (who) collaborated with brands and (Mona did/creators) too.

117.	Roger and Ralph started a water conservation initiative.	Last year (only/it was) Roger (who) organised water-saving campaigns and (Ralph did/workshops) too.
118.	Carol and Trish teach Kindergarten.	Last year (only/it was) Carol (who) planned educational activities and (Trish did/classroom routines) too.
119.	Juliet and Anne manage a makeup artistry business.	Last Monday (only/it was) Juliet (who) organised lipsticks and (Anne did/client appointments) too.
120.	Marco and Luca coach a youth cricket team.	Last week (only/it was) Marco (who) led training drills and (Luca did/the practice) too.

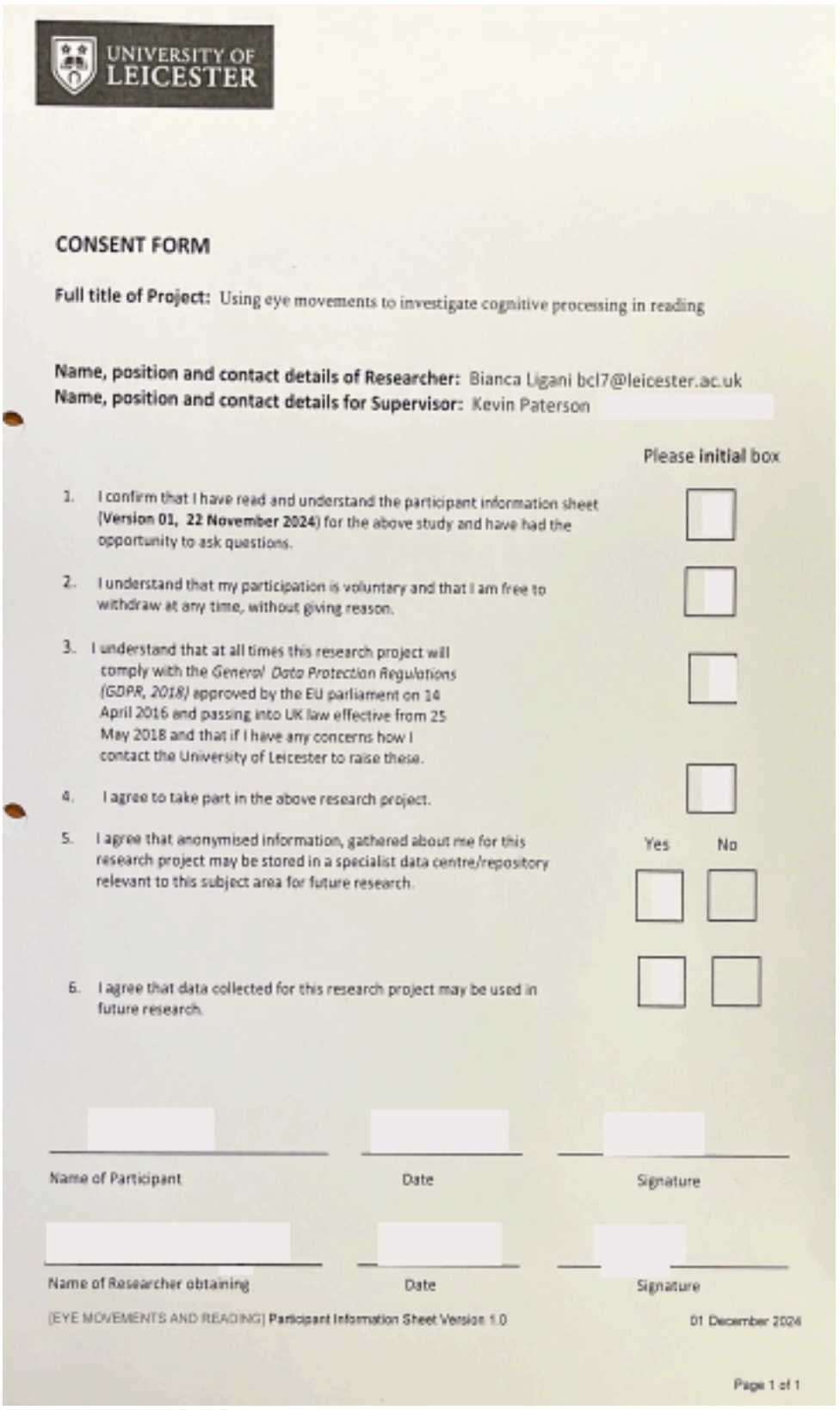
APPENDIX E – EXPERIMENTAL ITEMS OF THE EYE-TRACKING TASK IN BP

	Context sentence	Target sentence
1.	Marina e Olívia administram uma padaria em Belo Horizonte.	Ontem (só/foi) a Marina (quem) preparou café e (a Olívia também/chá também).
2.	Ricardo e Gustavo dão aula em uma escola em Juiz de Fora.	Na semana passada (só/foi) o Ricardo (quem) passou dever de casa e (o Gustavo também/projetos também).
3.	Bianca e Patrícia conhecem diferentes cidades no mundo.	No mês passado (só/foi) a Bianca (quem) visitou museus e (a Patrícia também/parques também).
4.	Gabriel e Rogério apresentam um podcast sobre viagens.	No último episódio (só/foi) o Gabriel (quem) entrevistou um guia e (o Rogério também/um agente também).
5.	Roberto e Matheus criam sites incríveis para clientes.	Hoje de manhã (só/foi) o Roberto (quem) criou um *layout* e (o Matheus também/gráficos também).
6.	Samuel e Marcelo trabalham em um hortifruti no Centro.	Semana passada (só/foi) o Samuel (quem) separou as frutas e (o Marcelo também/os legumes também).
7.	Jéssica e Renata têm uma loja de roupas em um shopping.	Ontem (só/foi) a Jéssica (quem) organizou novas peças e (a Renata também/as prateleiras também).
8.	Sofia e Aline produzem filmes independentes em Contagem.	No final de semana (só/foi) a Sofia (quem) filmou cenas e (a Aline também/paisagens também).
9.	Daniel e Otávio administram uma livraria aconchegante na cidade.	Hoje de manhã (só/foi) o Daniel (quem) limpou os livros e (o Otávio também/as mesas também).
10.	Carina e Elisa dão aulas comunitárias de yoga no parque.	Hoje (só/foi) a Carina (quem) ensinou alongamentos e (a Elisa também/exercícios também).
11.	Fernanda e Isabel pintam belos quadros sob encomenda.	Ontem (só/foi) a Fernanda (quem) pintou uma paisagem e (a Isabel também/um retrato também).

12. Rafael e Felipe são voluntários em um abrigo para animais.	Hoje (só/foi) o Rafael (quem) alimentou os cachorros e (o Felipe também/os gatos também).
13. Antônio e Osvaldo têm uma empresa focada em inovação.	Essa semana (só/foi) o Antônio (quem) conduziu reuniões e (o Osvaldo também/parcerias também).
14. Letícia e Natália coreografam danças para apresentações escolares.	Neste ano (só/foi) a Letícia (quem) ensaiou o balé e (a Natália também/a quadrilha também).
15. Camila e Larissa escrevem um blog sobre sustentabilidade.	Na última matéria (só/foi) a Camila (quem) listou produtos sustentáveis e (a Larissa também/dicas também).
16. Leandro e Murilo trabalham em uma academia da cidade.	Nessa semana (só/foi) o Leandro (quem) coordenou o spinning e (o Murilo também/o pilates também).
17. Ulisses e Evandro têm uma padaria no bairro deles.	Hoje (só/foi) o Ulisses (quem) assou os pães e (o Evandro também/os bolos também).
18. Simone e Vanessa organizam eventos comunitários em Betim.	Mês passado (só/foi) a Simone (quem) planejou um festival e (a Vanessa também/uma corrida também).
19. Vinícius e Guilherme criam aplicativos de celular para empresas.	Ano passado (só/foi) o Vinícius (quem) testou performance e (o Guilherme também/usabilidade também).
20. Lorena e Mirela mantêm um canal de comida no YouTube.	Semana passada (só/foi) a Lorena (quem) postou uma receita e (a Mirela também/uma crítica também).
21. Henrique e Caetano produzem um podcast sobre saúde mental.	No último programa (só/foi) o Henrique (quem) compartilhou dicas e (o Caetano também/histórias também).
22. Raíssa e Luana ensinam dança do ventre em uma academia.	Mês passado (só/foi) a Raíssa (quem) coreografou a apresentação e (a Luana também/as aulas também).
23. Francisco e Rodrigo criam conteúdo sobre tecnologia para a internet.	Esse mês (só/foi) o Francisco (quem) apresentou inovações e (o Rodrigo também/empresas também).

24. Benjamin e Tiago gerenciam um estúdio especializado em fotos.	Essa semana (só/foi) o Benjamin (quem) produziu um editorial e (o Tiago também/um ensaio também).
25. Nicole e Bárbara ministram cursos de arte para crianças.	No último sábado (só/foi) a Nicole (quem) preparou materiais e (a Bárbara também/atividades também).
26. Fabrício e Danilo gerenciam uma agência de viagem na cidade.	Hoje (só/foi) o Fabrício (quem) reservou voos e (o Danilo também/hotéis também).
27. Gilberto e Maurício treinam um time de futebol local.	No último treino (só/foi) o Gilberto (quem) conduziu o aquecimento e (o Maurício também/os coletivos também).
28. Beatriz e Amanda têm um perfil sobre maquiagem no Instagram.	Na última semana (só/foi) a Beatriz (quem) comprou um contorno e (a Amanda também/um batom também).
29. Clarissa e Elaine confeitam tortas e bolos para uma doceria.	Ontem (só/foi) a Clarissa (quem) decorou as tortas e (a Elaine também/os bolos também).
30. Samara e Joana arrecadam fundos para causas locais.	Ano passado (só/foi) a Samara (quem) captou dinheiro e (a Joana também/doações também).
31. Rodolfo e Nícolas prestam um serviço de lava-a-jato móvel.	Ontem (só/foi) o Rodolfo (quem) lavou automóveis e (o Nícolas também/ônibus também).
32. Alberto e Leonel adestram cachorros domésticos.	Na última sessão (só/foi) o Alberto (quem) ensinou comandos e (o Leonel também/truques também).

APPENDIX F – CONSENT FORM OF THE EYE-TRACKING TASK IN ENGLISH



UNIVERSITY OF LEICESTER

CONSENT FORM

Full title of Project: Using eye movements to investigate cognitive processing in reading

Name, position and contact details of Researcher: Bianca Ligani bcl7@leicester.ac.uk
Name, position and contact details for Supervisor: Kevin Paterson

Please initial box

1. I confirm that I have read and understand the participant information sheet (**Version 01, 22 November 2024**) for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason. ☐
3. I understand that at all times this research project will comply with the *General Data Protection Regulations (GDPR, 2018)* approved by the EU parliament on 14 April 2016 and passing into UK law effective from 25 May 2018 and that if I have any concerns how I contact the University of Leicester to raise these. ☐
4. I agree to take part in the above research project. ☐
5. I agree that anonymised information, gathered about me for this research project may be stored in a specialist data centre/repository relevant to this subject area for future research. Yes ☐ No ☐
6. I agree that data collected for this research project may be used in future research. ☐ ☐

Name of Participant _____ Date _____ Signature _____

Name of Researcher obtaining _____ Date _____ Signature _____

[EYE MOVEMENTS AND READING] **Participant Information Sheet Version 1.0** 01 December 2024

Page 1 of 1

APPENDIX G – TCLE OF THE EYE-TRACKING TASK IN BP

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Gostaríamos de convidar você a participar como voluntário (a) da pesquisa **PSICOLINGUÍSTICA E MULTIMODALIDADE: O PROCESSAMENTO DA LINGUAGEM EM TEXTOS MULTIMODAIS**. O motivo que nos leva a realizar esta pesquisa é entender a relação entre informações visuais ou informações presentes no contexto de interação e a linguagem verbal. Nesta pesquisa pretendemos investigar como se dá o acesso a informações verbais e visuais no processamento da linguagem.

Caso você concorde em participar, vamos fazer a seguinte atividade com você: **um teste de leitura em que você terá o movimento dos olhos mapeado na tela de um computador com a ajuda de um equipamento de rastreamento ocular**. Esta pesquisa tem alguns riscos, que são: (i) algum desconforto físico pela posição (sentada) e/ou tempo de exposição aos itens nas atividades experimentais; (ii) possibilidade de sentir algum desconforto ou constrangimento na leitura dos itens experimentais pelo seu conteúdo, (iii) possibilidade de se sentir constrangido no ato de realizar a tarefa experimental com o aparelho de rastreamento; (iv) possibilidade de identificação não intencional dos participantes da pesquisa. Mas, para diminuir a chance desses riscos acontecerem, vamos fazer intervalos de descanso quando o tempo da atividade experimental for superior a 20 minutos, os itens experimentais foram cuidadosamente elaborados e não abordam conteúdos sensíveis, o experimento é realizado individualmente, sem contagem de tempo de resposta e você vai receber todos os esclarecimentos necessários antes do início da tarefa, e, por fim, vamos registrar a sua participação com um código alfanumérico para que nenhum dado pessoal, nem mesmo o seu primeiro nome, seja divulgado.

Para participar deste estudo você não vai ter nenhum custo, nem receberá qualquer vantagem financeira. Apesar disso, se você tiver algum dano por causa das atividades que fizermos com você nesta pesquisa, você tem direito a buscar indenização. Você terá todas as informações que quiser sobre esta pesquisa e estará livre para participar ou recusar-se a participar. Mesmo que você queira participar agora, você pode voltar atrás ou parar de participar a qualquer momento. A sua participação é voluntária e o fato de não querer participar não vai trazer qualquer penalidade ou mudança na forma em que você é atendido(a). O pesquisador não vai divulgar seu nome. Os resultados da pesquisa estarão à sua disposição quando finalizada. Seu nome ou o material que indique sua participação não será liberado sem a sua permissão. Você não será identificado(a) em nenhuma publicação que possa resultar.

Este termo de consentimento encontra-se impresso em duas vias originais, sendo que uma será arquivada pelo pesquisador responsável e a outra será fornecida a você. Os dados coletados na pesquisa ficarão arquivados com o pesquisador responsável por um período de 5 (cinco) anos. Decorrido este tempo, o pesquisador avaliará os documentos para a sua destinação final, de acordo com a legislação vigente. Os pesquisadores tratarão a sua identidade com padrões profissionais de sigilo, atendendo a legislação brasileira (Resolução Nº 466/12 do Conselho Nacional de Saúde), utilizando as informações somente para os fins acadêmicos e científicos.

Declaro que concordo em participar da pesquisa e que me foi dada à oportunidade de ler e esclarecer as minhas dúvidas.

Belo Horizonte, __ __ de _____ de 2025

Assinatura do Participante

Assinatura do (a) Pesquisador (a)

Pesquisadoras responsáveis: Profa. Aline Alves Fonseca e Bianca Lignani

E-mail:

O CEP avalia protocolos de pesquisa que envolve seres humanos, realizando um trabalho cooperativo que visa, especialmente, à proteção dos participantes de pesquisa do
Brasil. **Em caso de dúvidas, com respeito aos aspectos éticos desta pesquisa, você poderá consultar:**
CEP - Comitê de Ética em Pesquisa com Seres Humanos - UFJF
Campus Universitário da UFJF
Pró-Reitoria de Pós-Graduação e Pesquisa
CEP: 36036-900
Fone: (32) 2102- 3788 / E-mail: cep.propp@ufjf.br

1