

Publication status: Not informed by the submitting author

THEMATIC RELATIONS: A STUDY ON CONCEPTUAL COMPOSITION IN BRAZILIAN PORTUGUESE

Mayda Rangel Gomes Peres, Marije Soto

<https://doi.org/10.1590/SciELOPreprints.7935>

Submitted on: 2024-01-22

Posted on: 2024-01-29 (version 1)

(YYYY-MM-DD)

The moderation of this preprint received the endorsement of:

Aniela Improta Franca (ORCID: <https://orcid.org/0000-0002-3614-2970>)

RESEARCH REPORT

THEMATIC RELATIONS: A STUDY ON CONCEPTUAL COMPOSITION IN BRAZILIAN PORTUGUESE

RELAÇÕES TEMÁTICAS: UM ESTUDO SOBRE COMPOSIÇÃO CONCEITUAL NO PORTUGUÊS BRASILEIRO

Mayda Rangel Gomes PERES (UFRJ/CAPES)

Master's student of the Postgraduate Program in Linguistics – Universidade Federal do Rio de Janeiro (UFRJ)

ORCID: <https://orcid.org/0000-0001-7675-4798>

Marije SOTO (UFRJ)

Professor of the Postgraduate Program in Linguistics – Universidade Federal do Rio de Janeiro (UFRJ)

ORCID: <https://orcid.org/0000-0003-4232-265X>

ABSTRACT

This study explores the processing of noun-prep-noun combinations in prepositional phrases in Brazilian Portuguese, focusing on the syntactic-semantic relationship marked by prepositions, implies the semantic sense in which the noun is being modified. We chose to contrast particularly two types of thematic relationships like 'made of' and 'made for', categorized as: material and function (e.g. 'toalha de crochê', crocheted cloth vs. agulha de crochê, crochet needle). To do so, we conducted two psycholinguistic experiments: Experiment 1 consisted of a self-paced reading/picture matching task to investigate the impact of semantic relationship on processing time; and Experiment 2 employed a cloze task (e.g. 'toalha de _____') and Likert scale likelihood judgments to norm the stimuli and assess cognitive preferences for the types of relation explored (i.e., material or function). Our results revealed slower reading times for the function type relation, suggesting a higher processing demand, possibly due to covert semantic computations required

for this relation. Cloze completions and likelihood ratings also indicated a cognitive preference for this type of relationship. We attribute these observed patterns to the robust syntactic-semantic relationship inherent to functional type relations, coupled with a broader range of potential response options. In contrast, material type relations demonstrated a higher level of convergence among participants for cloze responses, likely due to more restricted response options and, consequently, increased probability. Future EEG studies are proposed to identify specific moments in the cognitive processing of the constituents of these structures and assess possible anticipatory effects of prepositional phrase modifiers. These discoveries make a meaningful contribution to the broader literature on language processing, aligning with analogous patterns observed in different languages with a cognitive distinction between thematic relations and feature-based (i.e., telic and constitutive relations), such as in Wisniewsky & Love, 1998; Estes et al., 2003, 2011; Pustejovsky et al., 2016; Flick et al., 2021.

RESUMO

Este estudo explora o processamento de combinações substantivo-preposição-substantivo em sintagmas preposicionados no Português Brasileiro, com foco na relação sintático-semântica marcada por preposições, que implica o sentido semântico em que o substantivo está sendo modificado. Optamos por contrastar particularmente dois tipos de relações temáticas, como 'feita de' e 'feita para', categorizadas como: material e função (por exemplo, 'toalha de crochê', pano de crochê vs. 'agulha de crochê', agulha de crochê). Para isso, conduzimos dois experimentos psicolinguísticos: o Experimento 1 consistiu em uma tarefa de leitura automonitorada/picture matching para investigar o impacto da relação semântica no tempo de processamento; e o Experimento 2 empregou uma tarefa de Cloze (por exemplo, 'toalha de ____') e avaliações de probabilidade com escala Likert para normalizar os estímulos e avaliar preferências cognitivas para os tipos de relação explorados (ou seja, material ou função). Nossos resultados revelaram tempos de leitura mais lentos para o tipo de relação funcional, sugerindo uma demanda de processamento mais alta, possivelmente devido a computações semânticas encobertas necessárias para essa relação. O teste de Cloze e a escala de Likert também indicaram uma preferência cognitiva por esse tipo de relação. Atribuímos esses padrões observados à relação sintático-semântica robusta inerente às relações do tipo funcional, juntamente com uma gama mais ampla de opções de resposta potenciais. Em contraste, relações do tipo material demonstraram um nível mais alto de convergência entre os participantes para respostas de Cloze, provavelmente devido a opções de resposta mais restritas e, conseqüentemente, maior probabilidade.

Propomos futuros estudos de EEG para identificar momentos específicos no processamento cognitivo dos constituintes dessas estruturas e avaliar possíveis efeitos antecipatórios dos modificadores nos sintagmas preposicionados. Essas descobertas contribuem significativamente para a literatura mais ampla sobre o processamento da linguagem, alinhando-se a padrões análogos observados em diferentes idiomas, com uma distinção cognitiva entre relações temáticas e relações de características (i.e., relações télicas e constitutivas), como em Wisniewsky & Love, 1998; Estes et al., 2003, 2011; Pustejovsky et al., 2016; Flick et al., 2021.

KEYWORDS: Conceptual Combination. Thematic Relations. Psycholinguistic.

PALAVRAS-CHAVE: Combinação Conceitual. Relações Temáticas. Psicolinguística.

INTRODUCTION

Human thought and language possess an extraordinary ability to combine simple structures into more complex semantic structures (Martin & Baggio, 2019). At every instance when we produce language, we gather and combine words. This process is natural and enables us to create new and infinite sentences through compositions, yielding novel interpretations and concepts on the fly.

Examples of this process are modifier and noun combinations, when we join adjectives and nouns, (e.g. red wine) nouns and nouns (e.g. wine bottle), etc. Flick et al. (2021) observe that often the interpretation of these combinations require speakers to deduce thematic and semantic relations that are not explicitly stated in the combined expression, but rather are the results of conceptual associations primarily denoted by the semantics inherent to the words. Thus, a contrast in meaning may arise by the feature knowledge of constituent concepts (e.g. material or color) and the type of thematic relations they evoke (e.g. 'made for'). As an example, the authors cite the interpretation of material vs. function, in expressions such as 'trophy cabinet', as a place where trophies are kept, and 'metal cabinet', not as a place where metals are stored, but rather with an

interpretation of attribute, a cabinet made of metal.

A series of Magnetoencefalography (MEG) studies were show that heightened neurophysiological activation related to combinatory processes can be observed approximately 200-250 ms following the initiation of the phrasal head within its combinatorial context (for a review see Pyllkanen et al, 2019) in the left anterior temporal lobe (IATL), while different thematic relations modulate activation 100ms after phrasal head onset in the left posterior temporal lobe (IPTL). Thus, dissociating these types of conceptual combinations at a neurological level.

Although a growing body of studies have now mapped out combinatorial processes in English (Pyllkanen et al., 2019; Westerlund *et al.*, 2015; Flick et al., 2021; Fló et al., 2020; Zhang et al.,2015), not much is known about other languages, such as Portuguese, that may display other productive combinatory patterns. One obvious difference is that head-modifier order is different (i.e. 'white car' vs. 'carro branco'), which might impact combinatory expectancy. Nonetheless, in an Electroencefalography (EEG) study in Spanish, Fló et al. (2020) found similar sensitivity to combined stimuli in a similar time window as Pyllkanen et al. (2019) who found combinatory activation 250 ms after noun presentation in similar pairs in English. However, they did show that the experimental design (i.e. how predictable combined stimuli are) affect neurophysiological responses as evidenced by a pre-combinatorial effect.

Another difference is the fact that modification by prepositional phrase (PP) is a very frequent combinatory mechanism in Portuguese, so much so that, differently from English, material interpretations tend to be expressed by PPs rather than adjectives (e.g. 'metal cabinet' vs. 'armário de metal'). Thematic relations in BP have been studied through priming paradigms and ERP methodology, by comparing thematic pairs (e.g. *capacete-MOTO*, 'helmet-MOTORCYCLE') to associative pairs (e.g. *carro-MOTO*, 'car-MOTORCYCLE'), and by comparing the effect of directionality on thematic pairs (e.g. *CASCA-banana* vs. *BANANA-casca*, 'PEEL-banana' vs. 'BANANA-peel') (Gomes, 2010; Soto, 2014). Both studies found processing benefits for thematic priming, with faster latencies and lower N400 amplitudes. However, these studies did not control for the type

of thematic relationship, and did not present words in a grammatically natural syntactic context such as a prepositional phrase.

In this study, we intend to study processing of noun-noun combinations in prepositional phrases. In Portuguese, thematic relationships such as 'made of' and 'made for' are explicitly marked, primarily by the function of the preposition in the prepositional phrase (e.g. *toalha de crochê*, 'crocheted cloth' vs. *agulha de crochê*, 'crochet needle'). That is, there is a syntactic relationship that implies the semantic sense in which the noun is being modified (e.g. 'made of crochet' vs. 'made for crochet').

As the prepositional phrase can imply thematic relations, we chose to contrast two types of relation: material and function. The material relation expresses a 'made of' relationship, directly relating the name to the material it is composed of, e.g. *molho de vinho* (sauce made of wine). In addition, it seems of the various meanings and functions that can be expressed by the preposition 'de', its more semantically transparent meaning is associated with the material aspect, which might imply faster processing or a more salient response for material compared to function. Meanwhile, the functional relation expresses a 'made for' situation, requiring the speaker to covertly invoke a verb that would semantically express the functionality, e.g. *rótulo de vinho*, as a label made for labeling wine. In terms of processing cost, we may thus imagine function type relations to be more demanding than material type relations since they require an extra interpretative step and are also less predictable (i.e. more diverse).

We conducted two psycholinguistic experiments involving self-paced reading/picture matching and cloze/acceptability measures to investigate noun + PP combinations in Brazilian Portuguese (BP) and their respective processing costs, as well as to verify the normative aspects of stimuli for a future EEG experiment.

Norming stimuli is of great importance for empirical validity, but it can be challenging. This is because the creation of effective stimuli takes into account factors that are known to impact processing (e.g. frequency data, bigram frequency, or transition probability, etc.) that, in the context of Brazilian Portuguese, are not abundantly available in the accessible corpora. Given this, this work takes into consideration the need to control

normative aspects to mitigate these limitations, aiming to refine the stimuli for better control of the results. One way to assess probability is by applying a cloze test, in which participants are asked to complete a blank space in sentence/phrase frame (e.g. 'agulha de _____') and the proportional number of participants that opted for a given word stands for its probability (e.g. 20% for 'crochê' would show low to intermediate probability). A likelihood judgment might provide an indirect measure of the improbability of a target suggested by the experimenters on a Likert scale of 1 to 5, thus complementing cloze measures (e.g. an average of 4 for 'crochê' would indicate that while 'crochê' is not highly probable, it is not improbable). The ultimate goal of these tests is to make sure likelihood is equally distributed among stimuli and that it is not the defining factor in experimental effects. Moreover, this work aimed to investigate the effect of semantic relationships on processing using a novel paradigm, with self-paced reading combined with picture matching, allowing a preliminary insight into the processing of the studied conditions (i.e., material vs. function). Finally, besides obtaining results from cloze and likelihood experiments as a means of regulating the data, we aim to explore these data as well, so as to investigate cognitive preferences for either one of the relation types (i.e., material or function).

1.1. THEMATIC/SEMANTIC RELATIONS AND BEHAVIORAL EVIDENCE

To contextualize the studies presented here, the theoretical framework that constitutes the foundation for this work will be outlined. In pursuit of this objective, the literature will be examined with the aim of understanding the authors' perspectives on conceptual combination, thematic and semantic relations, as well as presenting empirical data from behavioral and EEG/ERP studies.

In order to explore how thematic relations influence processing, we chose to observe two types of relation: material, expressing a 'made of' connection where the noun directly relates to the material it is composed of, and function, expressing a 'made for' connection with the suppression of the verb that would explicitly indicate functionality.

These relations can be represented by what Pustejovsky and Jezek (2015) conceptualize as *Qualia*, which are mechanisms for the identification and differentiation of distinct properties in the representation of the core meaning of words, especially those associated with objects. Specifically, our analysis focuses on two main categories of *Qualia*: constitutive, which aims to encode information related to the materials that make up an object and the parts that constitute it (e.g. porcelain plate, i.e. made of porcelain); and telic, which aims to encode information concerning the function and purpose of the object in question (e.g. porcelain cabinet, i.e. made for storing porcelain).

With this relationship established, in the perspective of semantic analysis, Estes et al. (2003; 2011) make a fundamental distinction between thematic relations and attributive relations. In this sense, thematic relations are established when there is a temporal, spatial, causal, or functional relation between things that perform complementary roles in the same scenario or event' (Estes et al., 2003, p. 250). A manifestation of this phenomenon is evident in the interrelation between 'chair' and its semantic association to 'to sit', or 'pen' to 'to write', elucidating the intrinsic intertwining of the thematic relation with the functional aspect of the object, where the chair is employed for the act of sitting, and the pen for writing. Thus, the establishment of thematic relations reinforces the coherence between these concepts, given that pens are inherently associated with the action of writing.

On the other hand, attributive relations involve the attribution of properties to a noun. In this context, the term 'chair' may be associated with the property 'wood', indicating the material from which the chair is constructed (Estes, 2011, p. 275). These attributive relations highlight characteristics or qualities associated with the concept, contrasting with thematic relations that are linked to functions.

Under this view, attributive relations are not exactly like thematic relations (or perhaps a subtype), nor taxonomic relations (i.e. chair – furniture), but more similar to what Wisniewsky and Love (1998) call property relations. In earlier work, these authors use the term property relations to describe interpretations in which the head of a noun-noun combination is modified by one or more of the properties of the modifier noun that are

applicable to the set of properties of the head noun. In a series of experiments with novel noun-noun combinations they investigated whether there is a general tendency for speakers to assign either thematic or property relations (e.g. the property interpretation of a 'robin hawk' might point to a hawk with similar physical features as a robin).

Their findings showed that on the whole participants tended to prefer thematic relation interpretations; however, the similarity of nouns in a combination seemed to favor property relations (e.g. 'car truck' is interpreted as a truck having properties of a car), and a tendency for property reading increased - but thematic readings did not - when people were primed for it, thus revealing some degree of strategic control. This result was also confirmed by Estes and Jones (2009) who showed that participants consistently recognized a target word faster after a thematically related prime word (i.e. integrative priming), irrespective of the predictability of the priming type (they varied the proportion of related pairs between experimental lists). This led the authors to conclude that thematic relations are evoked automatically and not very amenable to strategic control.

Wisniewsky and Love (1998) also assessed a predominance for thematic relation types by extracting over a thousand existing noun-noun combinations for noun referents (mostly animals, plants, and artifacts) from reference manuals. Thus, in their view, while both mechanisms are equally relevant cognitive strategies for meaning inference, thematic relations are relatively more frequent.

Behavioral studies comparing the effects of the difference between these relations on processing cost and time show that in noun-noun combinations thematic relations are generally recognized faster and more easily than feature-based or taxonomic relations (for a review see Estes and Jones (2009)). Similar results were found in BP by Gomes (2010) and Soto (2014).

Contrary to these findings, there is a view that supports the idea of a 'thematic-to-taxonomic shift' and the 'syntagmatic-paradigmatic shift' (Estes; Golonka; Jones, 2011, P. 273), which claims that while children prefer thematic relation, adults prefer taxonomic relations. However, the empirical basis for this supposed shift is questionable, given that there is evidence that attributive preference in adults is strictly task dependent

and that older adults prefer thematic readings. These observations validate the idea that taxonomic and property relations are more influenced by strategic processing, while underscoring the significance of thematic relations in the construction of meanings and the comprehension of concepts, irrespective of any apparent preference for attributive relations.

In order to add further granularity to the understanding of thematic relations, we propose to follow Pustejovsky's approach by comparing preferences and processing costs associated to relations as mediated by telic and constitutive qualia. Recent neurophysiological studies have also been able to deepen the understanding of these relations by mapping out the timeline of the combinatory processes that underlie the interpretation of thematically bound concepts. Therefore, before presenting the experiments of the current study, we briefly review the results of the most relevant neurolinguistic studies.

1.2. CONCEPTUAL COMBINATIONS AND NEUROPHYSIOLOGICAL EVIDENCE

The combinatorial effect is fundamental for linguistic processing, but becomes complex as an object of analysis due to the possibility of different cognitive mechanisms that may underlie compositionality. Pylkkänen (2019) identifies at least three: (i) syntax; (ii) semantics (i.e., in the phrase 'red boat' it is logically possible that a boat has color properties); and (iii) conceptual (i.e., 'trophy cabinet' combines the concept of a cabinet with trophies, resulting in the concept of a cabinet for trophies).

Regarding the possibility of capturing these mechanisms, Pylkkänen (2019, p. 1) explains, for example, that in an experimental context, syntactic effects are challenging to distinguish from semantic effects because in natural language, syntactic changes often alter the meaning of the expression. Meanwhile, considering that syntax is difficult to vary while keeping meaning constant, the reverse is easier: it is possible to maintain the syntactic structures of experimental stimuli constant while varying the meaning.

Following this experimental rationale, Flick (2021) delved into semantic relations and minimal phrase composition by analyzing variations in thematic relations and characteristics modifications using MEG. He demonstrates that, in English, the neurophysiological responses imply a conceptual relationship primarily marked by the semantics of the individual word. For example, in the case of 'metal cabinet', (armário de metal) a relationship can be established directly since the metal trait is recognized as a potential attribute of a cabinet, resulting in a 'mixed' concept (it is a cabinet and it is made of metal). On the other hand, the conceptual combination of 'trophy cabinet' (armário para troféus) involves the implicit insertion of semantic-syntactic material, clarifying its relationship (e.g. a cabinet [for storing] trophies). In alignment with Pylkkanen's studies with adjective-noun combinations, the author demonstrates that the level of compositionality modulates negativity at 250 ms subsequent to the presentation of the head noun. However, only in the case of the trophy-cabinet type stimuli, resulted in additional activation in the left posterior temporal lobe. According to Flick (2021, p. 5131) these findings revealed that conceptual composition extends beyond a mere word combination; instead, phrase interpretation is grounded in implicitly expressed semantic relations.

In addition to studies using MEG, there are others that employed the Event Related Potential (ERP)/EEG technique, such as Zhang's research (2013), which hypothesized there are different levels of compositionality in various Chinese idiomatic expressions (e.g. low, medium, and high compositionality), varying across dimensions such as familiarity, literality, compositionality, and context. Similar to the MEG studies, the study demonstrated that the level of compositionality modulates negativity at 250 ms after the presentation of the target. The authors applied a semantic priming paradigm, in which short phrases expressing the literal meaning of idioms (i.e. to kick a bucket) were presented as primes, and idioms (i.e. kick the bucket - meaning 'to die') as targets. Amplitudes and peak latencies in response to target processing increased as the level of compositionality decreased (Zhang et al, 2013, p. 103). The authors attributed this effect to the degree of competition between activation in response to the literal meanings and

figurative meanings of the elements that composed the idioms. Competition was considered to be stronger, and to require more cognitive effort, when these two meanings are more distinct (i.e. when the idiom is less compositional). Thus, it seems that the compositionality addressed by Flick and Zhang is inherently an effect of conceptual structure (Flick et al, 2021; Zhang et al, 2013).

It is also important to consider possible differences between processing effects in light of differences in languages concerning word order, specifically that of head noun and modifier. In a recent study on composition using EEG, Fló et al.(2020) adapted a study by Bemis and Pykkänen (2011) with adjectives and nouns (e.g. 'red boat') to Spanish (e.g. 'bote rojo'), aiming to separate effects resulting from true composition from other more strategic processes related to expectation (Fló et al, 2020).

Flo et al. found a pre-combinatory effect (i.e. activation anticipating the presentation of the adjective), which they attributed to stimuli presentation effects: in a block design all stimuli were combinatory, which enhanced expectancy, and, consequently, the pre-combinatory effect. The authors preferred to explain this effect as task related and not word order (i.e. language specific) related, for which they present a series of arguments. First of all, there is evidence from a MEG study with Arabic, which also presents noun-modifier order similar to Spanish and BP, with similar activation-timing patterns as for English (Westerlund *et al.*, 2015). Moreover, similar activation has been observed for reversed pairs in English (i.e. 'boat red') in a study which required participants to combine the words conceptually in order to complete a picture-matching task (Bemis; Pykkänen, 2013). Finally, the authors suggest that pre-combinatory activation has been found with regular English adjective-noun pairs.

All in all, more studies are required, with a wider variety of combinatorial structures and semantic relation types to deepen the understanding of both task specific and language specific processing effects. One such contribution is the one proposed in the current study with post-nominal modification via PP. Not only might different thematic relations affect processing as foreseen by Flick et al. (2021), but also the presence of the preposition - a phrase projecting category - might enhance expectancy of combinatory

processes. Moreover, compared to the functional relation type, in the material relation type, the preposition 'de' takes on a more semantically transparent meaning which might also result in faster processing or a more salient response for material compared to function.

Up to this point, within the context of the literature on compositionality, neurophysiological studies in the neuroscience of language have been detailed, featuring contributions from Flick et al. (2021), Fló et al. (2020), and Zhang et al. (2015). The results of these studies align affirmatively with the hypothesis that there is a compositionality effect in the combination of components within a linguistic context.

1.3. THE PRESENT WORK - QUESTIONS AND HYPOTHESIS

These stimuli have been prepared for experimentation attempting to control for factors delineated by Zhang et al. (2015), with a primary focus on avoiding influences from production frequencies of stimuli, as well as acceptability and likelihood. Furthermore, the study aims to explore the effects of language factors, such as specific factors in Brazilian Portuguese, as constituent order and the presence of the preposition with a crucial role in syntactic and semantic mediation, drawing insights from the research of Fló et al. (2020). Additionally, the investigation seeks to elucidate the distinctions between thematic and semantic relations of material and function in terms of response time and processing, as described by Flick et al. (2021).

Experiment 1 aims to investigate the effect of semantic relationships (i.e., material vs. function) on processing in terms of timing and cognitive load by applying a self-paced reading task combined with picture matching. It aims to verify the hypothesis that FUN type relations involve implicit semantic computation, requiring an additional step for processing. Thus, we foresee that there is a slower processing time for this condition. Moreover, we hypothesize that the presence of the preposition serves as an explicit structural marker of combinatory processes. The semantically more transparent meaning of 'de' associated with the material aspect is expected to result in faster processing for

material compared to function type relations.

Additionally, for Experiment 2, we ran a cloze task with phrasal frames (e.g. *agulha de ____*) and asked likelihood judgments on proposed targets (e.g. *agulha de crochê*), using a 5 point Likert scale. The data from these tasks will be explored two-fold: (i) as a means of norming the data for a future EEG study, (ii) to observe whether there are any cognitive preferences for either one of the relation types (i.e., material or function).

In terms of norming, both the cloze task and the likelihood task may indicate whether proposed targets in the self-paced reading task are probable and likely to the participants. The cloze task may result in a cloze percentage for the chosen target if participants converge to a degree with our proposed target, whereas the likelihood measure indicates how likely a participant judges to have thought of a given target irrespective of whether the target was among the words suggested by the cloze measure. For the purposes of studying combinatory processes comparing the two relation types, stimuli must ideally not be too familiar (so as to avoid the risk of lexicalized or figurative combinations, such as *coração de pedra*, 'heart of stone'), nor be implausible (so as to avoid semantic violation). Additionally, ratings must be equally distributed between conditions. The constraints of the experimental design (e.g. the repetition of target words between conditions, number of syllables, frequency, etc) demand a delicate balance for which both more open ended inquiries (such as cloze) as well as more closed judgements (i.e. of an a priori proposed target) can be very informative.

Besides stimuli norming, we might also explore cloze and likelihood measures to observe cognitive preferences. In accordance with the literature (Estes et al., 2003, 2011; Pustejovsky et al., 2015; Wisniewsky and Love, 1998), we presume that there is a difference in cloze response choices for MAT and FUN, which raises the question of whether there is a preference for one of them. To this end, we might verify if the thematic interpretation (henceforth referred to as class) of the cloze response corresponds to the class of the target word selected by us. That is, if the cloze responses differ from the target, do they also diverge in class (i.e., FUN or MAT) or are they within the same class? A tendency towards FUN might be taken as indication of cognitive preference.

Considering this, another hypothesis is that MAT differs from FUN also due to comparatively restricted possibilities (i.e. there are a limited number of materials to be expressed or to apply to a given head noun), with the possibility that FUN categories generate more varied cloze responses and thus fewer repetitions (i.e. lower mean cloze percentage).

Finally, in terms of likelihood, we might observe a cognitive preference for the FUN condition if participants consistently judge these combinations to be more plausible than MAT conditions. That is because while telic (i.e. FUN) relations are expected to require more effort in establishing a thematic connection due to a form of complement coercion, they could also potentially lead to a more pronounced semantic-syntactic connection. The outcome of this might be increased processing demands, which Estes (2011, p. 275) suggests there is an intensified relational aspect in the inferred meaning, and it may also be reflected in a preference for FUN relations in cloze tasks and likelihood judgements.

2. PARTICIPANTS

91 university students participated in the experiment (69 females; average age: 23.05), with no history of language problems and normal or corrected-to-normal vision. The research was previously approved by the Ethics Committee of Instituto de Estudos e Saúde Coletiva da Universidade Federal do Rio de Janeiro / IESC - UFRJ under the number CAEE 76522623.5.0000.5286. All participants signed a Consent Form.

3. EXPERIMENTAL DESIGN AND STIMULI

The two experiments shared a similar design by introducing the variable semantic relationship type, presenting two conditions: functional (FUN) and material (MAT). The stimuli consisted of a noun (repeated between conditions) in prepositional phrase, functioning as a modifier of a head noun that would vary in accordance to the condition. Thus, following the design: one for function (modifier specifies the function of the head

noun, considered a thematic relationship; e.g. 'garrafa de leite' ('bottle of milk'), and one for material (modifier specifies the material of the head noun, considered a feature-based relationship; 'pudim de leite' ('milk pudding')), with a fixed prepositional phrase mediated by 'de' connected to its name that could relate to both material or function interpretation (See Figure 1).

Stimuli were created in order to provide a certain level of control over word and combination frequencies. They were generated based on combinations found in the Brazilian Corpus (Portuguese Web 2020 (ptTenTen20), more specifically the subcorpus Brazilian TLD.br) using the Sketch Engine program.

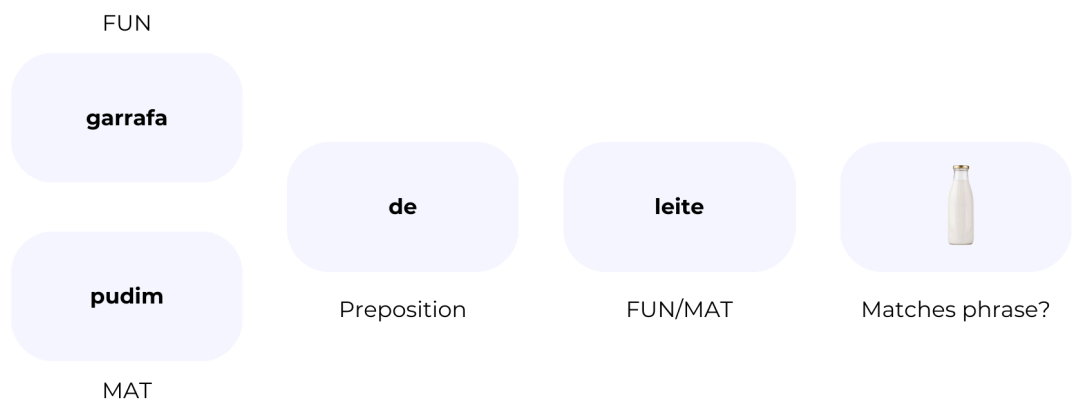
This program allowed us to control measures such as Cooccurrences (the total number of occurrences of the collocate within the selected left and/or right range), Candidates (the total number of occurrences of the collocate in the whole corpus, or subcorpus if selected) and LogDice, in relation to the collocation of the main word. The point was to capture the sequence or combination of words that occur together more often than would be expected by chance, on the one hand, and to avoid highly frequent combinations, on the other (See online repository for list: <https://osf.io/yx9zu/>).

Due to the semantic and combinatory restrictions, individual frequencies and word length varied. Zipf scores for each of the words were extracted from the *Léxico do Português* (ESTIVALET, 2020) platform. Although head noun frequencies varied (from 0.477 to 4.1304), there was no significant difference between the average frequencies of each condition ($t = 0.099$, $df = 19$, $p\text{-value} = 0.922$). However, we would like to note that this corpus, as it is written-based, tends to underestimate frequencies of words for concrete everyday objects. The nouns within the modifier PPs presented frequencies ranging from 1.415 to 5.277, but there was no difference between conditions since these stimuli are repeated across conditions. The number of letters for the head noun varied from 4 to 10 letters, but there was no significant difference between conditions ($t = 1.828$, $df = 19$, $p\text{-value} = 0.0837$).

For Experiment 1, each participant read 20 stimuli, 10 in each condition (FUN and MAT), along with 10 distractor stimuli. For Experiment 2, each participant also read 20

stimuli, 10 in each condition (FUN and MAT), along with 10 distractor stimuli. The stimuli were randomly divided into 4 lists, in such a way that for each list each stimulus was presented in only one of the tasks, ensuring that the same list was not repeated for the same participant.

Figure 1 - Stimulus presentation



Font: The Authors (2023)

However, it is crucial to clarify that the Brazilian Portuguese corpora do not provide all the necessary measures, such as bigram frequency or transition probability. This creates slight challenges in meticulously controlling the stimuli, demanding alternative approaches such as the psycholinguistics experiments of Experiment 2 described below.

4. PROCEDURE

The complete experimental design consisted of two separate experiments to probe the processing of material vs. function, with (i) self-paced reading and picture matching task and (ii) cloze test and acceptability judgment (Likert scale) of stimuli.

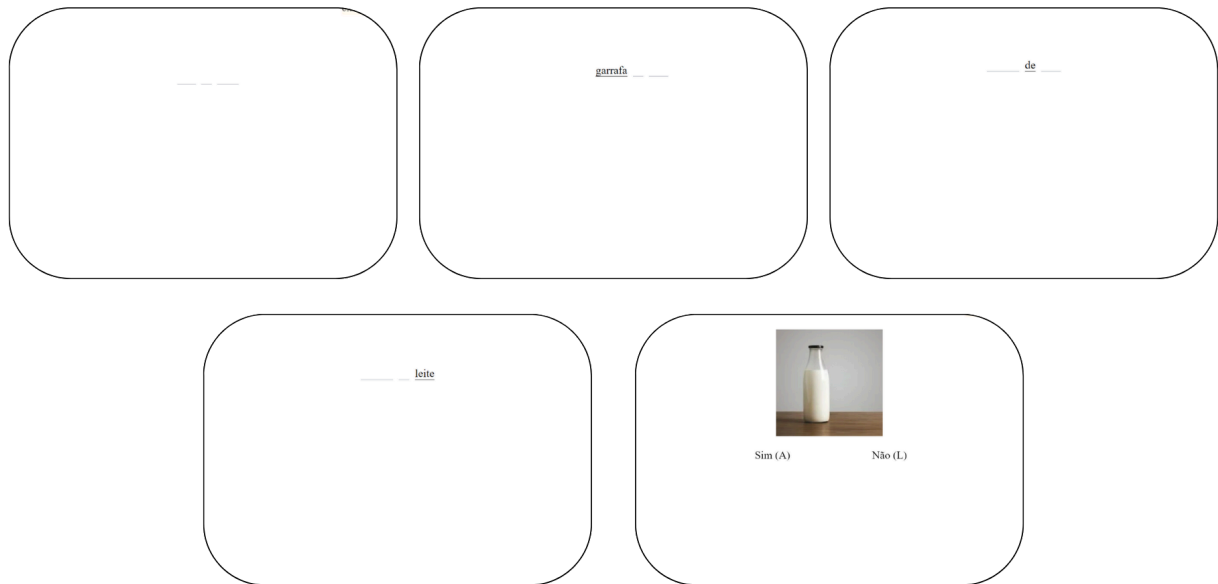
All tests were conducted online via the PC IBEX platform (ZEHR.; SCHWARZ, 2018). The test was administered remotely, and participants were instructed to perform it

in a silent environment, preferably on a computer.

4.1. EXPERIMENT 1

Our aim was to measure the processing costs associated with phrases featuring thematic relations of function material. At the beginning of each stimulus, only the spaces corresponding to the parts of the stimulus appeared on the screen (See Figure 2). To read the first word of each stimulus, the participant pressed a button, which revealed it. Pressing the button again removed the read part, while revealing the subsequent word. After the last word, an image appeared for which the participant had to respond whether it matched the previous phrase. For example, in the case of 'vaso de vidro' ('vase of glass'), the image could be an image of a vase made of glass or a random object. The person could choose between two options on the screen: YES or NO, by pressing specific buttons that were informed during the training phase. The questions varied, sometimes focusing on the combination, sometimes on the material, reducing the chances of the participant realizing what was happening. The picture matching task served to check participants' attention, to make sure participants engaged in conceptual combination, while drawing participants' attention.

Figure 2 - Self-paced reading



Source: The Authors (2024)

4.2. EXPERIMENT 2

For the Cloze and acceptability test, participants were presented with a phrasal frame with a gap to complete with what they judged with the first word that came to mind, for example, 'copo de ___' ('glass of ___'). Participants were indirectly induced to respond with only one word in the training session. During the test, they responded by typing it in (see Figure 3).

Immediately afterward each stimuli, a declared measure of target probability followed, using a 5 point Likert scale. Participants were asked to rate from 1 to 5 how likely they would have been to think of the suggested word (i.e. our target word) as an alternative option to the word they themselves had suggested, ranging from 1 (unlikely) to 5 (very likely). Before starting the test, a training session was conducted to ensure that the participant understood the tasks.

The stimuli was distributed in a Latin square design to ensure that all participants see all conditions, but no participant sees the same sentence in different conditions. Thus,

four versions of the same test were created.

Figure 3 - Cloze test

Complete a sequência com a primeira palavra que você pensou.

pastel de ____

Aperte enter para enviar sua resposta.

Em uma escala de 1 a 5, qual a probabilidade de você pensar nessa palavra como uma outra opção de preenchimento?

pastel de pizza

Nenhuma	Média	Muita		
1	2	3	4	5

Source: The Authors (2024)

5. ANALYSIS

In Experiment 1, two dependent variables were recorded: reaction time (RT) and accuracy. Data organization and statistical analysis were done in RStudio (version 2021.09.1+372, Rstudio TEAM, 2021). For both measures, generalized mixed models were used with semantic relation type (funcion x material) as fixed effect, with participants and items as random factors. For accuracy measures, a binomial distribution from the lme4 package (Bates et al. 2015) was applied. One participant was eliminated due to low accuracy (<50%). We extracted RTs for the preposition and the noun in the modifier PP phrase, which we considered to be the critical area of analysis. RT distribution approached gamma distribution, which was used, also from the lme4 package (Lo; Andrews, 2015). Values below 100ms and above 2500 ms were eliminated based on cognitive relevance for the task at hand (Smith; Levy, 2013; Kapteijns, Hintz, 2021).

For Experiment 2, we applied glme models with binomial distribution to test for target agreement in the cloze test, while applying a cumulative link mixed model for ordinal data from the Ordinal package in order to analyze the effect of semantic relation type on the Likert-scale data from the likelihood test (Endresen; Janda, 2017).

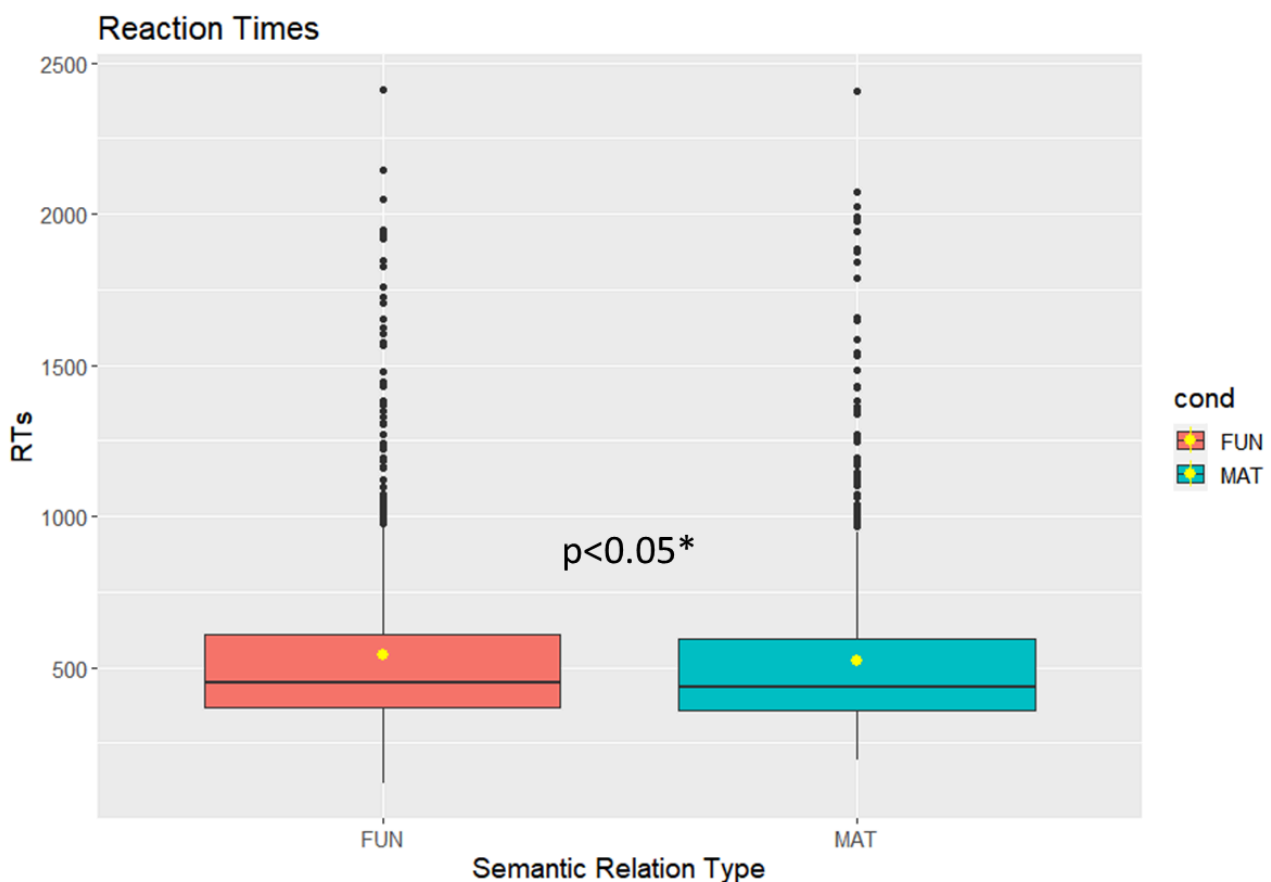
Scripts can be found among the supplementary materials: <https://osf.io/yx9zu/>

6. RESULTS

6.1. EXPERIMENT 1

There was an overall effect ($X^2(1) = 4.106, p < 0.05^*$) for the semantic relation type observed, with faster RTs for MAT. The average reaction time (RT) for FUN was 541 ms (SD: 292 ms); whereas for MAT, it was 521 ms (SD: 273 ms) (see Graph 1).

Graph 1 - Reaction Times



Source: The Authors (2024)

The accuracy values were 98.9% for FUN and 98.7% for MAT, indicating a high

number of correct responses in both cases, without significant difference ($X^2(2) = 0.316$, $p = 0.85$). There was also no difference in response times for this task ($X^2(1)=0.364$, $p=0.547$).

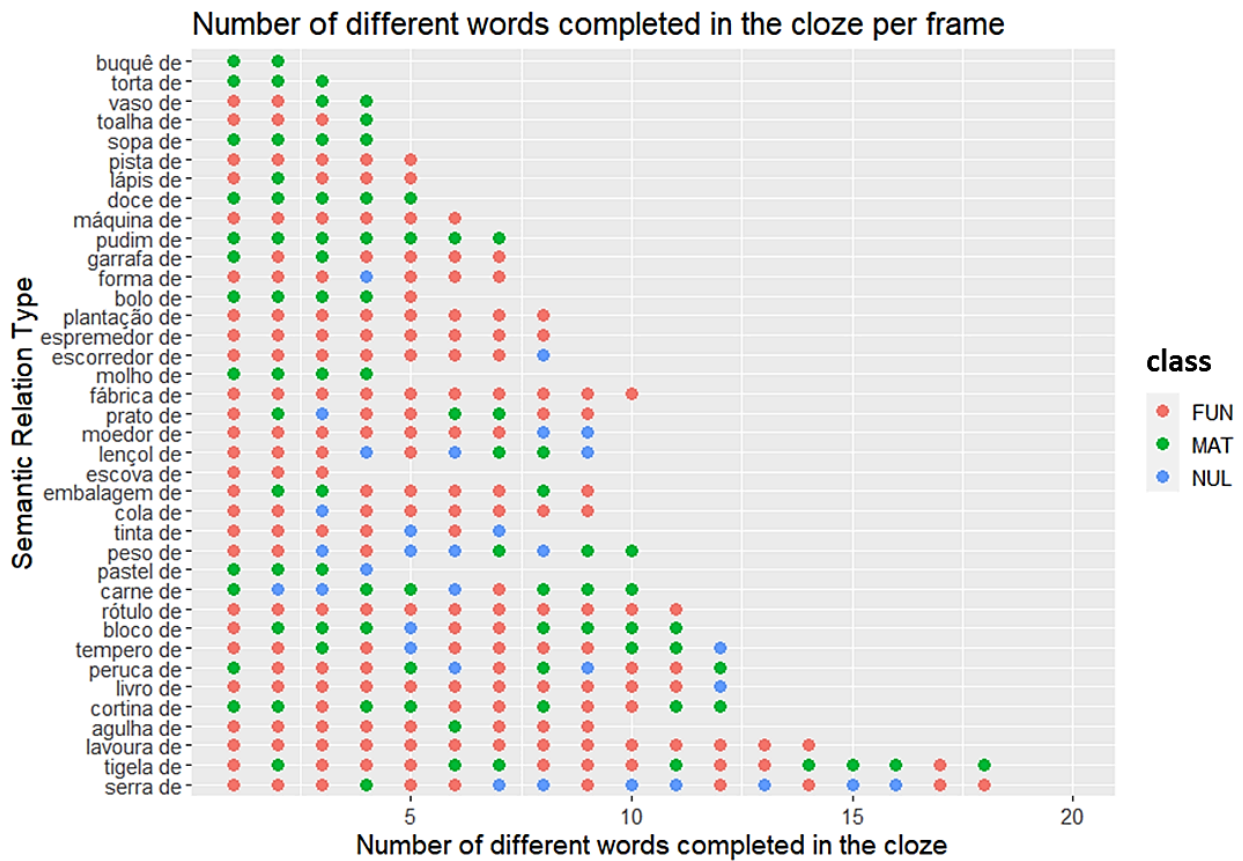
6.2. EXPERIMENT 2

6.2.1 COGNITIVE PREFERENCE FOR RELATION TYPE

The Cloze test indicated general preference, given that 61% of cloze responses were of the FUN type, compared to 34.2% of the MAT type, and 4.8% unclassified (e.g., with the possibility of polysemic interpretation, as in 'serra de...' which could mean tool or mountainous region). Comparing the divergence from the target semantic relation (of the stimulus as presented in Exp.1) between conditions, there was a significant difference ($X^2(1) = 99.78$, $p < 0.001^{***}$) in participants' responses regarding their preference for MATERIAL or FUNCTION. In the target semantic relation of FUN, for the most part, the relationship composed of the words filled in the cloze was also characterized as FUN, with 84%, compared to 10.7% for MAT. For the predicted MAT relationship, the words filled in the cloze that also formed a MAT relationship accounted for 57.8%, but a substantial portion was related to FUN, with 38%.

If we look at the variety of words by semantic relation type (average number of different words completed for each frame), the result of the Cloze task showed a significant difference ($F(2)=5,84$, $p<0.01^{**}$), with the average number of words for FUN being 10.3 and for MAT being 9.28. This shows that, although diversity varies from item to item, on average, cloze items establishing a FUN relationship tend to yield more varied responses.

Graph 2 - Number of different words completed in the cloze per frame



Source: The Authors (2024)

This also means that FUN presents lower average cloze values, given the probability of Cloze responses being distributed across a greater number of candidates. On average, the max. cloze percentages (for the most completed word) for FUN were 42.0%, compared to 56.6%. This greater diversity for FUN, on the one hand, and greater restriction for MAT, on the other is also reflected in the number of times words were repeated in each semantic relation type. FUN had 31 (out of 153) words used more than once (2 to 5 repetitions); whereas MAT had 19 (out of 62) words used more than once (2 to 5 repetitions), indicating that words establishing a MAT relationship are more likely to be repeated, with 30.65%, compared to FUN with 20.26%.

Regarding likelihood (Likert scale), the average and distribution of responses showed an overall effect of relationship type on likelihood judgments ($X^2(1) = 31.873, p < 0.001^{***}$), with overall higher percentages for 3 (medium) to 5 (very likely) for FUN,

whereas percentages were relatively higher for MAT for 1 (unlikely) and 2 (See Table 1).

Table 1 - Likelihood in Likert responses

How likely..?	LIKERT	FUN (%)	MAT (%)
unlikely	1	28.00	36.44
	2	12.89	14.00
medium	3	21.78	20.00
	4	15.56	10.67
very likely	5	21.78	18.89
(% judgment per level)			

Source: The Authors (2024)

6.2.2 STIMULI NORMING

There was relatively low convergence between participants' cloze responses and the expected target for both conditions, with only 6.5% (FUN) and 3.3% (MAT) similar answers given ($X^2(1)=0.509$, $p=0.476$). We also observed the cloze probability for these cases of convergence, that is, the proportion of participants that chose a given word for a specific frame matching our target word (e.g. for 'agulha de ____', 20% chose the word 'crochê'). For MAT, cloze probability was an average of 40.9% (average deviation: 18.31%). For FUN, this was 23.4% (average deviation 6.95%). However, the average for MAT has a higher average deviation with more disparity between high and low values.

Overall probability judgment averages varied, from 1.05 (for 'máquina de cimento', 'cement machine') to 4.50 (for 'pudim de leite', milk pudding). Comparing MAT and FUN conditions overall, there was no statistical difference ($X^2(1)=1.77$, $p=0.18$). Nonetheless, we specifically looked at how comparable probability judgments on the Likert scale were for

two phrases with the same target word. Differences between judgments per item were 1.02 on average, varying from 0.35 (for 'algodão', in 'plantação de algodão', with 2.15, and 'lençol de algodão', with 2.50; 'cotton planting' and 'cotton sheet') to a maximum of 2.25 (for 'leite', in 'garrafa de leite', with 2,25, and 'pudim de leite', with 4,50; 'milk bottle' and 'milk pudding') (see Table 2 for all judgment measures).

Table 2 - Likelihood per item/stimuli

alvo	FUN	med_FUN	MAT	med_MAT	DIF
algodão	algodão _ plantação de algodão	2,15	algodão _ lençol de algodão	2,50	0,35
cabelo	cabelo _ escova de cabelo	3,59	cabelo _ peruca de cabelo	4,04	0,44
café	café _ moedor de café	3,32	café _ doce de café	1,50	1,82
chocolate	chocolate _ fábrica de chocolate	3,95	chocolate _ torta de chocolate	3,35	0,6
cimento	cimento _ máquina de cimento	1,05	cimento _ bloco de cimento	2,75	1,7
crochê	crochê _ agulha de crochê	3,90	crochê _ toalha de crochê	1,80	2,1
feijão	feijão _ tempero de feijão	2,80	feijão _ sopa de feijão	2,15	0,65
flor	flor _ vaso de flor	4,04	flor _ buquê de flor	4,41	0,37
gelo	gelo _ forma de gelo	2,75	gelo _ pista de gelo	2,30	0,45
isopor	isopor _ cola de isopor	3,11	isopor _ embalagem de isopor	1,91	1,2
laranja	laranja _ espremedor de laranja	4,05	laranja _ bolo de laranja	3,10	0,95
leite	leite _ garrafa de leite	2,25	leite _ pudim de leite	4,50	2,25
louça	louça _ escorredor de louça	3,50	louça _ prato de louça	2,25	1,25
madeira	madeira _ serra de madeira	3,32	madeira _ lápis de madeira	1,64	1,69
papel	papel _ peso de papel	2,90	papel _ livro de papel	1,75	1,15
pizza	pizza _ molho de pizza	2,75	pizza _ pastel de pizza	2,30	0,45
soja	soja _ lavoura de soja	1,91	soja _ carne de soja	2,71	0,81
tecido	tecido _ tinta de tecido	2,82	tecido _ cortina de tecido	2,41	0,41
vidro	vidro _ fábrica de vidro	1,68	vidro _ tigela de vidro	2,93	1,25
vinho	vinho _ rótulo de vinho	2,10	vinho _ molho de vinho	1,50	0,6

Source: The Authors (2024)

7. DISCUSSION

In this work, the self-paced reading task of Experiment 1 yielded results that align with our hypothesis, indicating there are processing differences between thematic and featured-based relations. Slower reaction times confirmed that there is a tendency for

higher processing cost associated with thematic relations of function. As presumed, this result could be explained due to the need to process covert semantic relations, as in a verb after the preposition, as in 'agulha de V crochê,' with the verb 'fazer' filling the gap (see section 1.1). Such additional computations possibly demand higher cognitive load, explaining the increased processing time.

Conversely, for the material aspect, we propose that the relationship of the prepositional phrase already explicitly directs the inferred meaning, not requiring the mediation of the verb. In this sense, the relationship between the noun and the prepositional phrase occurs directly, possibly because the most transparent meaning of the preposition 'de' is associated with a feature-based relation before function. Therefore, it is likely that the direct relationship between the noun (e.g., 'garrafa', 'bottle') and the prepositional phrase (e.g., 'de vidro', as 'of glass') induces faster processing for material-type relations compared to functional ones.

Nonetheless, we must concede that the difference in processing time, albeit statistically significant, was very subtle, which leads us to caution, especially in light of the methodological difficulties concerning the control of stimuli properties. Future experiments might present a full latin square design such that not only the modifying noun is repeated among conditions, but the head noun as well. In this experiment, we opted not to apply this design due to the difficult balance between well-formed and logical combinations and perfect experimental control. However, if we look at reading times for the first segment (the head noun), they vary, mostly under the effect of word length ($X^2(1)=131.79$, $p<0.001^{***}$). It is not entirely unimaginable that there might be a spill over of this effect on the following segments, thus introducing possible confounds to the main effect of semantic relation. We feel that the data from the cloze and likelihood test have the potential to improve stimulus control for future experiments in order to avoid such confounds. Furthermore, more fine-grained techniques such as EEG/ERP do not require the methodological artifact of self-paced reading, thus enabling better control of presentation rates and time-locking neurophysiological responses to the stimuli. In this way, it might be easier to understand the contribution of each of the constituents to

combinatory processing as the structure unfolds incrementally.

For picture matching responses, there was no significant result to differentiate any of the relations, as the accuracy rate was very high for both. This indicates a high level of attention in the test, but it does not show a relationship between the task difficulty and processing time.

Additionally, we conducted Experiment 2, which allowed us to verify post-processing inferences of thematic relations in participants' Cloze responses, while also contributing to the improvement and control of stimuli for future studies.

The tests provided support for our hypothesis, demonstrating that, in many cases, there is a difference in response choices between MAT and FUN, with a preference for FUN regarding all words in the cloze test. It is important to note that, in the cloze test, the phrasal frames presented to participants did not explicitly guided towards a given category (FUN or MAT). Even so, participants showed a clear preference for FUN for all words completed in the cloze (61%), and even the imagined target was directed towards MAT, in many cases participants still preferred FUN (38%). This is interesting because it confirms that people tend to first respond with the thematic relationship of function, especially in items with a previous sense of world knowledge and frequency of use: e.g. 'pencil of...' with the imagined target for MAT being 'wood', responses like 'coloring' were common, tending towards FUN. This is evidence of a stronger thematic-syntactic relationship between the constituents of the phrase due to the hidden semantic meaning (i.e., the verb interacts with the preposition, guiding the inferred functional aspect), grounding this conceptual relationship more deeply (see section 1.1). Thus, even though there is a greater cognitive effort to process (see Experiment 1), this relationship is strengthened, validating the trend of responses in this class.

Furthermore, it is worth considering that some responses did not fit into either one of the selected categories (i.e., MAT or FUN), which we then labeled as NUL. These cases encompassed some categories of 'content', such as 'bowl of...' with the target for 'glass', obtaining responses like 'soup', 'porridge' (for discussions on this type of relation and its implications for measure or count readings, see Partee; Bosrchev, 2012; Khrizman *Et Al.*,

2018). We also addressed polysemic categories that could have a double interpretation, such as 'form', which could be the object (i.e. a specifically shaped container) or the way of doing something.

It is interesting to highlight that the hypothesis that MAT differs from FUN also due to comparatively restricted possibilities in terms of post nominal modification is confirmed by the fact that there is a greater number (i.e. variety) of words used for FUN (153), compared to MAT (62). The material relationship might be more restrictive because there is a smaller list of possible materials that could constitute any object, e.g, metal, glass, stone, wood, etc., a range that becomes even more restricted within the confounds of the features of a given object (e.g. a towel by its very nature can only be made of some type of cloth). To name an example, for 'towel of' there was only one response designating MAT, namely, 'cloth', while for FUN, the responses obtained were 'table', 'face', 'bath', thus presenting much more variety within the same class of FUN.

So, although MAT has a smaller number/variety of responses, the cloze probability of target matches responses are higher (40.9%), indicating higher predictability by way of a clear restriction of possibilities. For FUN, they present a greater variety of possibilities; however, with lower cloze probability for target matches (23.4%). Also, if we look at average cloze probabilities for the most frequent words that were not target matches, MAT cloze probability (56.5%) is significantly higher than FUN (42.0%), further confirming higher predictability for MAT.

The second test of Experiment 2 observed the likelihood of items based on subjects' judgments, where they had to choose between unlikely, medium, or very likely. Interestingly, we observed a general effect of semantic relation, with less acceptability for MAT items - contrary to a possible higher predictability, despite a smaller range of possibilities (i.e., fewer possible words), there was no greater acceptability in the words chosen by us. This may have occurred because for each item the likelihood judgment immediately followed the cloze task. Thus, when items obtained cloze responses with tendencies towards FUN, subsequent likelihood judgments may have been biased in that direction, including for the phrasal frames with imagined target items corresponding to the

MAT condition, lowering their likelihood rating. In future studies, this might be avoided by separating these tasks into different blocks and by not repeating any items in both tasks.

In terms of ensuring engagement in combinatory processing, we aimed at creating stimuli that were at once likely to occur (i.e. plausible) but not too familiar (i.e. pre-established idiomatic expression). Flick *et al.* (2021) discusses the methodological issue of novel combinations vs. existing combinations. On the one hand, novel and unknown combinations are sure to engage combinatory processes (the example he gives is 'sponge memory', p. 3); on the other, they might also rouse implausibility effects or conscious monitoring (e.g. 'is this a thing?'), which are to be avoided. In our likelihood ratings we observed that for both classes, the highest percentages were for 'unlikely' (28.0% for FUN and 36.44% for MAT) and for the summed percentage for 'unlikely' to 'medium' (61.67% for FUN and 69,44% for MAT). This result is noteworthy because it reveals that the items are not perceived to be very likely (i.e. plausible), thus, we feel that on the whole the stimuli aided in eliminating a possible effect of lack of engagement, while also avoiding some very frequent pre-established idiomatic expressions.

Finally, in order to control for confounds between conditions (MAT and FUN), predictability and frequency between the two phrases containing the same item need to converge. We take both cloze probability and likelihood ratings to be an indirect measure of the predictability and frequency of a phrase in the absence of transition probability and bigram frequency indices in the available corpora for Brazilian Portuguese. Considering that completion responses are relatively unguided and free, probability measures serve as the most controlled indication of stimulus properties in this regard. Here, we observe that, although there is no interference of higher/lower frequency affecting participants' perception when comparing the two conditions globally (on average the difference between the judgment of two phrases with the same item is 1.02), if we look items individually, there is still some inconsistency. This underscores the need for a more careful selection among the stimuli, choosing to include only those with less variability in judgment between them, ensuring there is no disparity in final measures due to frequency differences in stimuli. Thus, a more empirically reliable result could be achieved. To

accomplish this, we intend to conduct additional tests of diverse nature (e.g., explicit plausibility assessments or requesting paraphrases, among others), for this aspect warrants further investigation.

Another methodological issue that can be enhanced to achieve more controlled results are some of the items in the 'FUN' and 'NUL' categories. Some of the nouns we used for the FUN condition, such as 'fábrica' (factory) might be considered deverbal nouns. These cases might engage a different computation, where, instead of the processing of a covert associated verb as we suggested for FUN, a verb might be activated by way of morphological decomposition of the noun itself. This might need to be reviewed.

Another example is when the semantic relationship can be interpreted as container or as a partitive ('garrafa de leite', bottle of milk). This leads us to consider two aspects regarding these phenomena: a) some cases might indeed be interpreted as partitive, attributing an alternative semantic interpretation; b) the morphological relation between noun and the verb, as in 'moedor de café' (coffee grinder) and fábrica, would still imply an additional effort to 'verbalize' and establish the thematic relation (leaving implicit task), consistent with the obtained result that the processing was slower.

Overall, our results tentatively confirmed our hypotheses of there being covert and, thus, more costly processing for FUN type relations as well as a cognitive preference for this type of relation, as is suggested in the literature also for other languages (Wisniewsky and Love, 1998, Estes et al., 2003, 2011; Pustejovsky et al., 2016; Flick *et al.*, 2021). We attribute this both due to the strong syntactic-semantic relationship and the broader range of response possibilities. Meanwhile, MAT presented a higher level of convergence in target matching responses, possibly to the restricted response options and therefore greater predictability. However, future EEG studies might be able to pinpoint at which moment combinatorial cognitive processing kicks in and whether specific attributes of the PP modifier in BP exert any anticipatory effect on these computations. It is worth noting that the effects of linguistic factors, such as specific features of Brazilian Portuguese, including constituent order and the presence of the preposition yielded results that were compatible with those obtained in previously mentioned studies in other languages and

other structure types (such as adjective noun and noun noun modification) (see Section 1.2). However, considering its high temporal resolution, the ERP technique, which we intend to apply in the future, might be a more suitable avenue for observing these online linguistic processes in greater detail.

BASIC INFORMATION

Competing Interests: The authors declare no competing interests.

Data Accessibility Statement: <https://osf.io/yx9zu/>

Contributor Roles (Contribuição de autoria): Mayda R. G. Peres: Conceptualization, Data curation, Investigation, Writing – original draft; Marije Soto: Methodology, Software, Supervision, Validation, Writing – review & editing.

Ethics and Consent: The research was previously approved by the Ethics Committee of Instituto de Estudos e Saúde Coletiva da Universidade Federal do Rio de Janeiro / IESC - UFRJ under the number CAEE 76522623.5.0000.5286. All participants signed a Consent Form.

Disclosure of Funding Sources: Self-funded; Master's scholarship from Capes.

CITATIONS

BATES, Douglas; MAECHLER, Martin; BOLKER, Ben; WALKER, Steven. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 2015.

BEMIS, Douglas K.; PYLKKÄNEN, Liina. Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of*

Neuroscience, v. 31, n. 8, p. 2801-2814, 2011.

BEMIS, D. K., & PYLKKÄNEN, L.. Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands, 2013, PLoS ONE, 8. <https://doi.org/10.1371/journal.pone.0073949>.

ENDRESEN, A.; JANDA, L. A. Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Research Design and Statistics in Linguistics and Communication Science*, v. 3, n.2, p. 217-250, 2017. DOI: <https://doi.org/10.1558/jrds.30822>.

ESTES, Zachary; GOLONKA, Sabrina; JONES, Lara L. Thematic thinking: The apprehension and consequences of thematic relations. In: *Psychology of learning and motivation*. Academic Press, 2011. p. 249-294;

ESTES, Zachary. A tale of two similarities: Comparison and integration in conceptual combination. *Cognitive science*, v. 27, n. 6, p. 911-921, 2003.

ESTES, Zachary; JONES, Lara L. Integrative priming occurs rapidly and uncontrollably during lexical processing. *Journal of Experimental Psychology: General*, v. 138, n. 1, p. 112, 2009.

ESTIVALET, G. L. (2020). Variáveis lexicais e ortográficas no acesso lexical das palavras do português brasileiro. *Linguística*, 16(1), 264-277. <https://doi.org/10.31513/linguistica.2020.v16n1a31539>

FLICK, Graham; ABDULLAH, Osama; PYLKKÄNEN, Liina. From letters to composed concepts: A magnetoencephalography study of reading. *Human brain mapping*, v. 42, n. 15, p. 5130-5153, 2021.

FLÓ, Emilia; CABANA, Álvaro; VALLE-LISBOA, Juan C. EEG signatures of elementary composition: Disentangling genuine composition and expectancy processes. *Brain and Language*, v. 209, p. 104837, 2020.

GOMES, J. N. Investigating The Syntax Semantics Interface: An Erp Study On The Dissociation Between N400 And P600 Effects. Postgraduate Certificate—UFRJ: 2010.

KAPTEIJNS, B.; HINTZ, F. Comparing predictors of sentence self-paced reading times: Syntactic complexity versus transitional probability metrics. *PLoS One*. Jul 12;16(7), 2021. <https://doi.org/10.1371/journal.pone.0254546>

KHRIZMAN, Keren et al. Portion readings are count readings, not measure readings. In: *Proceedings of the 20th Amsterdam colloquium*. Amsterdam: ILLC, 2015. p. 197-216.

KILGARRIFF, Adam; BAISA, Vít; BUŠTA, Jan; JAKUBÍČEK, Miloš; KOVÁŘ, Vojtěch; MICHELFEIT, Jan; RYCHLÝ, Pavel; SUCHOMEL, Vít. The Sketch Engine: ten years on. *Lexicography*, 1: 7-36, 2014.

KIM, Songhee; PYLKKÄNEN, Liina. How the conceptual specificity of individual words affects incremental sentence composition: MEG evidence. *Brain and Language*, v. 218, p. 104951, 2021.

KUZNETSOVA, Alexandra; BROCKHOFF, Per B.; CHRISTENSEN, Rune H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 2017.

LENTH, Russell V. lsmeans: Least-Squares Means. Rpackageversion2. <http://CRAN.R-project.org/package=lsmeans>, 2015.

LO, Steson; ANDREWS, Sally. To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6, 2015 <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01171/full>

MARTIN, Andrea E.; BAGGIO, Giosuè. Modelling meaning composition from formalism to mechanism. *Philosophical Transactions of the Royal Society B*, v. 375, n. 1791, p. 20190298, 2020.

PARTEE, Barbara H.; BORSCHEV, Vladimir. Sortal, relational, and functional interpretations of nouns and Russian container constructions. *Journal of Semantics*, v. 29, n. 4, p. 445-486, 2012.

PUSTEJOVSKY, James; JEZEK, Elisabetta. *Introducing Qualia Structure. A Guide to Generative Lexicon Theory*, 2015.

PYLKKÄNEN, Liina. The neural basis of combinatory syntax and semantics. *Science*, v. 366, n. 6461, p. 62-66, 2019.

PYLKKÄNEN, Liina. Neural basis of basic composition: what we have learned from the red-boat studies and their extensions. *Philosophical Transactions of the Royal Society B*, v. 375, n. 1791, p. 20190299, 2020.

RSTUDIO TEAM. *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA, 2021 URL <http://www.rstudio.com/>

SMITH, N. J.; LEVY, R. The effect of word predictability on reading time is logarithmic, *Cognition*, Volume 128, Issue 3, 2013, Pages 302-319, ISSN 0010-0277.

SOTO, Marije; FRANÇA, Aniela Improta. The functional analysis of the N400 component: lexical access, integration or can we have it both ways?. *Revista Linguística*, v. 16, n. Esp., p. 521-562, 2018.

WESTERLUND, Masha et al. The LATL as locus of composition: MEG evidence from English and Arabic. *Brain and Language*, v. 141, p. 124-134, 2015.

WISNIEWSKI, Edward J.; LOVE, Bradley C. Relations versus properties in conceptual combination. *Journal of memory and language*, v. 38, n. 2, p. 177-202, 1998.

ZHANG, Hui et al. ERP correlates of compositionality in Chinese idiom comprehension. *Journal of Neurolinguistics*, v. 26, n. 1, p. 89-112, 2013.

ZHANG, Linmin; PYLKKÄNEN, Liina. The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, v. 111, p. 228-240, 2015.

ZEHR, J.; SCHWARZ, F. PennController for Internet Based Experiments. (IBEX), 2018. Disponível em: <https://osf.io/md832/>. Acesso em: 22 de ago. de 2022.

This preprint was submitted under the following conditions:

- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a [Creative Commons CC-BY](#) license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.